

# Guide to nonlinear potential estimates

Tuomo Kuusi · Giuseppe Mingione

Received: 20 September 2013 / Accepted: 17 December 2013 / Published online: 9 January 2014  
© The Author(s) 2014. This article is published with open access at SpringerLink.com

**Abstract** One of the basic achievements in nonlinear potential theory is that the typical linear pointwise estimates via fundamental solutions find a precise analog in the case of nonlinear equations. We give a comprehensive account of this fact and prove new unifying families of potential estimates. We also describe new fine properties of solutions to measure data problems.

## Contents

|     |   |    |
|-----|---|----|
| 1   | A synopsis                                    | 2  |
| 2   | From linear to nonlinear                      | 4  |
| 3   | Basic nonlinear estimates                     | 8  |
| 4   | Universal potential estimates                 | 14 |
| 5   | Maximal-potential estimates                   | 17 |
| 5.1 | Non-endpoint maximal estimates                | 19 |
| 6   | Equations with coefficients                   | 20 |
| 7   | Interlude on measure data problems            | 21 |
| 7.1 | Potential estimates and fundamental solutions | 24 |
| 8   | Fine properties of solutions via potentials   | 25 |

---

Communicated by Neil Trudinger.

The authors are supported by the ERC grant 207573 “Vectorial Problems” and by the Academy of Finland project “Regularity theory for nonlinear parabolic partial differential equations”.

---

T. Kuusi

Aalto University Institute of Mathematics, P.O. Box 11100, 00076 Aalto, Finland  
e-mail: [tuomo.kuusi@aalto.fi](mailto:tuomo.kuusi@aalto.fi)

G. Mingione (✉)

Dipartimento di Matematica e Informatica, Università di Parma,  
Parco Area delle Scienze 53/a, Campus, 43124 Parma, Italy  
e-mail: [giuseppe.mingione@math.unipr.it](mailto:giuseppe.mingione@math.unipr.it)

|      |   |    |
|------|---|----|
| 9    | Regularity and corollaries                          | 26 |
| 10   | A basic comparison estimate                         | 29 |
| 11   | A sequence of comparison estimates                  | 33 |
| 12   | Proof of Theorems 8 and 10                          | 39 |
| 12.1 | A density property of $\alpha$ -harmonic functions  | 40 |
| 12.2 | Setting of the constants                            | 42 |
| 12.3 | Iterating quantities                                | 45 |
| 12.4 | Iteration chains                                    | 46 |
| 12.5 | Density and decay properties along iteration chains | 46 |
| 12.6 | Iteration and conclusion                            | 51 |
| 12.7 | Proposition 1 and Theorem 8                         | 54 |
| 13   | Proof of Theorem 12                                 | 57 |
| 13.1 | A VMO-type result                                   | 59 |
| 14   | Proof of Theorem 15                                 | 61 |
| 14.1 | Proof in the case of energy solutions: beginning    | 61 |
| 14.2 | Setting of the constants and a sequence             | 62 |
| 14.3 | Non-degenerate iteration chains and proof of (180)  | 64 |
| 14.4 | Very weak solutions                                 | 68 |
| 15   | Theorem 3 as a corollary                            | 70 |
| 16   | Proof of Theorem 9                                  | 71 |
| 16.1 | General setting                                     | 71 |
| 16.2 | Estimates on the excess                             | 72 |
| 16.3 | Pointwise estimate and Lebesgue point criterion     | 74 |
| 16.4 | Use of the maximal estimate and conclusion          | 75 |
| 17   | Proof of Theorem 16                                 | 76 |
| 17.1 | Smallness of the excess                             | 77 |
| 17.2 | Proof that $x$ is a Lebesgue point of $u$           | 79 |
| 18   | Back to the roots: Theorem 1 (sketch)               | 80 |

## 1 A synopsis

The aim of this paper is to give a rather comprehensive introduction to nonlinear potential estimates, i.e., pointwise estimates for solutions to quasilinear, possibly degenerate elliptic equations via linear and nonlinear potentials. The paper contains both new and old results. They fall into two categories. The first consists of those results that have been proved elsewhere, and that are here given in different and/or streamlined version. The second contains new results that are presented for the first time. These fill some of the gaps that were making the current theory still somewhat incomplete.

To ease the reading, the already known results will be stated together with the reference to the corresponding original paper where they have appeared for the first time. The new ones will be presented pointing at the place of this paper where the proof can be found. In general, the first part of the paper is devoted to the presentation of the general setting, with the statements of the main theorems; this goes up to Sect. 8. The remaining parts are instead devoted to the proofs.

We shall start from a presentation of the classical results valid for linear elliptic equations in Sect. 2, and this will serve to give the general guideline to the topics we are going to cover in the nonlinear case. The first potential estimates for nonlinear equations will be introduced in Sect. 3. There the classical pointwise estimates will be presented. In the subsequent Sects. 4 and 5 we shall instead give a class of estimates aimed at unifying the theory. These allow to frame the classical pointwise potential inequalities in a more general setting, allowing for estimates of both size and oscilla-

tions of solutions and their derivatives, including fractional derivatives. At this stage we shall use “intermediate” Wolff and Riesz potentials and various fractional maximal operators. Especially, we prove Theorem 10, whose role is to unify estimates for  $u$  and  $Du$  via a family of estimates for suitable fractional operators of  $u$ . We shall eventually add a few remarks on the case of equations with coefficients in Sect. 6. All the results up to Sect. 6 are presented for *energy solutions*, that is for solutions belonging to the natural energy space  $W^{1,p}$  associated to the problems considered. In Sect. 7 we then turn to the case of general measure data problems, that is when *very weak solutions* come into the play. This will give us the opportunity, in Sect. 8, to present a few interesting consequences of potential estimates. Specifically, we shall prove two theorems about the possibility of describing the fine behaviour of solutions to nonlinear measure data problems and of their gradient via potentials. These theorems are the nonlinear analogs of classical results about fine properties of solutions to linear equations, which are usually derived via linear potential; see Remark 3 below. In Sect. 9 we then gather a series of regularity results that follow as a corollary of the theory presented. Some of these results are well-known and they are now framed in the general and unifying context of nonlinear potential estimates. The remaining sections are devoted to the proofs of the results introduced in the preceding ones, and their titles are self-explaining. In this paper we are not going to deal with parabolic problems, for which we refer to [42–44].

Before starting, we find useful to establish some notation. In what follows we denote by  $c$  a general positive constant, possibly varying from line to line; special occurrences will be denoted by  $c_1, c_2, \bar{c}_1, \bar{c}_2$  or the like. All these constants will always be *larger or equal than one*; moreover relevant dependencies on parameters will be emphasized using parentheses, i.e.,  $c_1 \equiv c_1(n, p, \nu, L)$  means that  $c_1$  depends only on  $n, p, \nu, L$ . We denote by

$$B(x, r) \equiv B_r(x) := \{\tilde{x} \in \mathbb{R}^n : |\tilde{x} - x| < r\}$$

the open ball with center  $x$  and radius  $r > 0$ . When not important we shall omit denoting the center as follows:  $B_r \equiv B(x, r)$ . Moreover, with  $B$  being a generic ball with radius  $r$ , we will denote by  $\sigma B$  the ball concentric to  $B$  having radius  $\sigma r, \sigma > 0$ . Unless otherwise stated, different balls in the same context will have the same center. The symbol  $\Omega$  will denote a bounded open subset of  $\mathbb{R}^n$  and we shall always consider the case  $n \geq 2$ . With  $\mathcal{O} \subset \mathbb{R}^n$  being a measurable subset with positive measure, and with  $g: \mathcal{O} \rightarrow \mathbb{R}^n, n \geq 1$ , being a measurable map, we shall denote by

$$(g)_{\mathcal{O}} \equiv \int_{\mathcal{O}} g \, d\tilde{x} := \frac{1}{|\mathcal{O}|} \int_{\mathcal{O}} g(\tilde{x}) \, d\tilde{x}$$

its integral average; here  $|\mathcal{O}|$  denotes the Lebesgue measure of  $\mathcal{O}$ . The map  $g$  will be typically a gradient. We shall denote

$$|g|(\mathcal{O}) := \int_{\mathcal{O}} |g| \, d\tilde{x}$$

therefore adopting a unified notation for both measures and integrals of  $L^1$ -maps. In the following,  $\mu$  will always denote a Borel measure with finite total mass, which is

initially defined on a certain open subset  $\Omega \subset \mathbb{R}^n$ . Since this will not affect the rest, with no loss of generality, all such measures will be considered as defined in the whole  $\mathbb{R}^n$  so that  $|\mu|(\mathbb{R}^n) < \infty$ . We shall denote by  $\mathcal{M}_b$  the space of all such measures.

## 2 From linear to nonlinear

When considering linear elliptic equations, a powerful tool for establishing the qualitative properties of solutions is given by representation formulas via fundamental solutions. Eventually, these lead to consider linear Riesz potentials and singular integrals. Let us recall the situation for the simplest example, which is of course given by the classical Poisson equation

$$-\Delta u = \mu. \quad (1)$$

For the sake of exposition, we initially consider the last equation in the whole  $\mathbb{R}^n$ , with  $n \geq 2$ . Here  $\mu$ , which is eventually taken to be a measure, is again for simplicity assumed to be a smooth and compactly supported function, while  $u$  is the unique solution which decays to zero at infinity. The point we are interested in now is that  $u$  can be recovered via convolution with the so called fundamental solution

$$G(x, y) \approx \begin{cases} |x - y|^{2-n} & \text{if } n \geq 3 \\ -\log |x - y| & \text{if } n = 2 \end{cases} \quad (2)$$

and this means that the following representation formula holds:

$$u(x) = \int_{\mathbb{R}^n} G(x, y) d\mu(y). \quad (3)$$

The formula in the last display allows to shift the study of solutions to the analysis of a related integral operator. It is therefore time for the following:

**Definition 1** (*Riesz potentials*) Let  $\alpha \in (0, n]$ ; the linear operator defined by

$$I_\alpha(\mu)(x) := \int_{\mathbb{R}^n} \frac{d\mu(\tilde{x})}{|x - \tilde{x}|^{n-\alpha}}, \quad (4)$$

is called the  $\alpha$ -Riesz potential of  $\mu$ , where  $\mu$  is a Borel measure defined on  $\mathbb{R}^n$ .

Now, (3) allows to conclude with the following pointwise potential estimates:

$$|u(x)| \lesssim I_2(|\mu|)(x) \quad \text{and} \quad |Du(x)| \lesssim I_1(|\mu|)(x). \quad (5)$$

The first inequality actually holds in the case  $n \geq 3$ ; the second one follows differentiating (3). Since the behaviour of Riesz potentials with respect to various relevant function spaces is known (see for instance [1, 26]), using (5) allows to infer *size* properties of solutions from those of the relevant potentials, and in a sharp way. For instance,

the following regularizing property of Riesz potentials is known for every  $q > 1$  and  $\alpha > 0$  such that  $0 < \alpha q < n$ :

$$\|I_\alpha(\mu)\|_{L^{\frac{nq}{n-\alpha q}}(\mathbb{R}^n)} \leq c(n, \alpha, q) \|\mu\|_{L^q(\mathbb{R}^n)}.$$

As an immediate corollary, it follows for instance that  $\mu \in L^q$  implies that  $Du \in L^{nq/(n-q)}$  whenever  $1 < q < n$ . Moreover, estimates in further function spaces can be obtained as well, and these yield borderline regularity results. An instance is given by the following Lorentz spaces criterion:

$$\mu \in L(n, 1) \implies I_1(|\mu|) \in L^\infty \implies Du \in L^\infty. \quad (6)$$

We recall that the Lorentz space  $L(q, \gamma)(\Omega)$ , for  $1 \leq q < \infty$  and  $\gamma > 0$ , consists of those measurable maps  $\mu$  satisfying the condition

$$\int_0^\infty \left( \lambda^q |\{\tilde{x} \in \Omega : |\mu(\tilde{x})| > \lambda\}| \right)^{\gamma/q} \frac{d\lambda}{\lambda} < \infty \quad (7)$$

so that  $\mu \in L(n, 1)$  means in particular that

$$\int_0^\infty |\{\tilde{x} \in \Omega : |\mu(\tilde{x})| > \lambda\}|^{1/n} d\lambda < \infty. \quad (8)$$

A point, which is less emphasised concerning estimates (5) is that (3) also allows to bound *oscillations* of the solution  $u$  in terms of suitable Riesz potentials. This point actually turns out to be of primary interest for us here. To see this, by using the elementary inequality

$$\left| |x - \tilde{x}|^{2-n} - |y - \tilde{x}|^{2-n} \right| \leq c(n) \left| |x - \tilde{x}|^{2-n-\alpha} + |y - \tilde{x}|^{2-n-\alpha} \right| |x - y|^\alpha,$$

which is valid whenever  $x, y, \tilde{x} \in \mathbb{R}^n$ , we get the new representation formula

$$|u(x) - u(y)| \leq c \left[ I_{2-\alpha}(|\mu|)(x) + I_{2-\alpha}(|\mu|)(y) \right] |x - y|^\alpha, \quad 0 \leq \alpha \leq 1. \quad (9)$$

The previous estimate allows to infer, again via potentials, information on the solutions in function spaces measuring smoothness rather than size. For instance, the Hölder continuity criterion

$$I_{2-\alpha}(|\mu|) \in L^\infty \implies u \in C^{0,\alpha}$$

is immediate. Again, prescribing that  $I_{2-\alpha}(|\mu|) \in L^q$  implies that  $u$  belongs to the so called Calderón space  $C_q^\alpha$ ; see Definition 2 below, and [17, 40], for a discussion. In a similar way one gets estimates for the gradient

$$|Du(x) - Du(y)| \leq c [I_{1-\alpha}(|\mu|)(x) + I_{1-\alpha}(|\mu|)(y)] |x - y|^\alpha, \quad 0 \leq \alpha < 1. \quad (10)$$

As a matter of fact, formulas (9)–(10) can be read as nonlocal, fractional derivatives versions of the classical estimates in (5). This is clarified by the following definition:

**Definition 2** (*Calderón spaces*) Let  $\alpha \in (0, 1]$ ,  $q \geq 1$ , and let  $\Omega \subset \mathbb{R}^n$  be an open subset. A measurable function  $v$ , finite a.e. in  $\Omega$ , belongs to the Calderón space  $C_q^\alpha(\Omega)$  if and only if there exists a nonnegative function  $m \in L^q(\Omega)$  such that

$$|v(x) - v(y)| \leq [m(x) + m(y)] |x - y|^\alpha \quad (11)$$

holds for almost every couple  $(x, y) \in \Omega \times \Omega$ .

This is just another way to say that  $v$  has “fractional derivatives”. The advantage is that nonlocality is reduced to a minimal status: only two points are considered in (11). Moreover, when  $\Omega$  is a suitably regular domain, Calderón spaces are closely related to the usual fractional Sobolev spaces  $W^{\alpha,q}$ ; see [17]. The function  $m$  plays in fact the role of a fractional derivative of  $v$  of order  $\alpha$  in the  $L^q$ -sense. Definition 2 is implicit in the work of DeVore and Sharpley [17], where the authors fix the canonical choice  $m = M_\alpha^\#(v)$  that is indeed always possible in (11) (see Proposition 1 below). The symbol  $M_\alpha^\#(v)$  denotes the standard fractional sharp maximal operator introduced in the following (take  $\Omega = \mathbb{R}^n$  and  $R = \infty$ ):

**Definition 3** (*Fractional sharp maximal operator*) Let  $\beta \in [0, 1]$ ,  $x \in \Omega$  and  $R \leq \text{dist}(x, \partial\Omega)$ , and let  $f \in L^1(\Omega)$ ; the function defined by

$$M_{\beta,R}^\#(f)(x) := \sup_{0 < r \leq R} r^{-\beta} \int_{B(x,r)} |f - (f)_{B(x,r)}| d\tilde{x}$$

is called the restricted (centered) sharp fractional maximal function of  $f$ .

Notice that for  $\beta = 0$  the previous definition gives back the classical sharp maximal operator of Fefferman and Stein (in the restricted version), which is in fact denoted by  $M_R^\#(f)$ . Definition 2 and the interpretation of  $m$  as the  $\alpha$ -order fractional derivative of  $v$  allow to read the inequalities in (9)–(10) as a way to bound fractional derivatives of  $u$  via intermediate Riesz potentials. We may express this concept in the following imprecise yet suggestive way:

$$|\partial^\alpha u(x)| \lesssim I_{2-\alpha}(|\mu|)(x), \quad 0 \leq \alpha < 2. \quad (12)$$

Needless to say, that last formula has only a symbolic meaning. At this point it is clear that the classical estimates (5) embed in the family of interpolating estimates (12) as particular borderline cases corresponding to the choices  $\alpha = 0$  and  $\alpha = 1$ , respectively. In turn, the estimates in (12) provides a unifying approach to the regularity of solutions to (1). They allow to get results both in spaces aimed at measuring the size of functions and in those measuring smoothness.

One of the aims of this paper is to show that completely similar estimates actually hold for a large class of quasilinear, possibly degenerate equations. A potential theory which is completely analogous to the linear one can be constructed in the nonlinear case too.

### 3 Basic nonlinear estimates

The main question is now whether and in which sense estimates like (5) and (12) extend to the case of solutions to quasilinear elliptic equations of the type

$$-\operatorname{div} a(Du) = \mu \quad \text{in } \Omega. \quad (13)$$

In the most general case, the right hand side  $\mu$  is a Borel measure with finite total mass in  $\mathbb{R}^n$ , i.e.  $\mu \in \mathcal{M}_b$ . The one in (13) is a nonlinear version of the Poisson equation when the assumptions

$$\nu \operatorname{Id} \leq \partial a(z) \leq L \operatorname{Id} \quad (14)$$

are considered. The numbers  $0 < \nu \leq L$  provide bounds for the lowest and highest eigenvalue of the matrix  $\partial a(\cdot)$ , respectively ( $\nu = L = 1$  gives the case of the Laplacean operator). We also want to consider cases where Eq. (13) might be degenerate, thereby examining for instance the so called  $p$ -Laplacean operator:

$$-\operatorname{div} (|Du|^{p-2} Du) = \mu. \quad (15)$$

In order to catch the essential properties of the equation in (15), we shall consider general vector fields  $a: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . These are assumed to be  $C^1$ -regular and to satisfy the following *growth and ellipticity assumptions*:

$$\begin{cases} |a(z)| + |\partial a(z)|(|z|^2 + s^2)^{1/2} \leq L(|z|^2 + s^2)^{(p-1)/2} \\ \nu(|z|^2 + s^2)^{(p-2)/2} |\lambda|^2 \leq \langle \partial a(z) \lambda, \lambda \rangle \\ p \geq 2 \end{cases} \quad (16)$$

whenever  $z, \lambda \in \mathbb{R}^n$ . The case  $p = 2$  gives back (14). We remark that we confine ourselves to the case  $p \geq 2$  as we are mainly interested to present the main ideas in the most accessible way; for results in the case  $p \leq 2$  we refer to [22, 38, 40, 58]. In the rest of the paper  $\nu, L$  are fixed parameters whose role is to establish the rate of ellipticity of the vector field  $a(\cdot)$ . The role of  $s \geq 0$  is more peculiar. This parameter serves to distinguish the degenerate ( $s = 0$ ) from the nondegenerate one ( $s > 0$ ). In this last case a model is given, by for instance taking  $s = 1$ , by the nondegenerate equation  $p$ -Laplacean type equation (see [49])

$$-\operatorname{div} ((1 + |Du|^2)^{(p-2)/2} Du) = \mu.$$

Sometimes we shall also consider equations with measurable coefficients, that is when the vector field  $a(\cdot)$  exhibits an explicit dependence on the variable  $x$ , i.e.,

$$-\operatorname{div} a(x, Du) = \mu. \quad (17)$$

In this case, we shall use a set of assumptions which is weaker than (16), namely, we shall consider a Carathéodory vector field  $a: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  such that

$$\begin{cases} |a(x, z)| \leq L(|z|^2 + s^2)^{(p-1)/2} \\ \nu(|z_1|^2 + |z_2|^2 + s^2)^{(p-2)/2}|z_1 - z_2|^2 \leq \langle a(x, z_1) - a(x, z_2), z_1 - z_2 \rangle \\ p \geq 2 \end{cases} \quad (18)$$

are satisfied whenever  $z_1, z_2, x \in \Omega$ . These assumptions are, up to adjusting the constants  $\nu, L$  in a universal way, implied by those in (16). Results for the case when  $x \rightarrow a(x, \cdot)$  is more regular and satisfy (16) uniformly with respect to  $x \in \Omega$  are described in Sect. 6 below.

Our primary emphasis will be on a priori estimates. This means that, according to a scheme which is typical in regularity theory, we shall mainly confine ourselves to present results in form of a priori estimates for more regular solutions and data. We are therefore most of the times considering weak energy solutions to (17), that is functions  $u \in W^{1,p}(\Omega)$  satisfying

$$\int_{\Omega} \langle a(\tilde{x}, Du), D\varphi \rangle d\tilde{x} = \int_{\Omega} \varphi d\mu \quad (19)$$

for every choice of  $\varphi \in C^\infty(\Omega)$  with compact support in  $\Omega$ . Here we are still assuming that  $\mu$  is a Borel measure with finite mass, but we could assume that  $\mu \in C^\infty$  as well, stating the same results involving only the total variation of  $\mu$ , considered as a measure. Results for the original context would then follow via approximation. Such an approach is obviously restrictive when considering general measure data problems. Indeed, observe for instance that due to the monotonicity properties of the operator assumed in (16), considering a solution that belongs to the space  $W^{1,p}(\Omega)$  automatically implies that the measure  $\mu$  belongs to the dual  $W^{-1,p'}$ , which is certainly not the case for any measure when  $p \leq n$ . As a matter of fact, in general, distributional solutions to measure data problems do not belong to the natural energy space  $W^{1,p}(\Omega)$  and for this reasons they are called *very weak solutions*. Treating the case of general measures needs then greater care, already in specifying the notion of solution one is dealing with. We are briefly discussing these aspects in Sects. 7 and 14.4 below, where we shall see how to extend the results presented for energy solutions to the case of general solutions to measure data problems.

Since we are going to deal with local results, we shall use in a standard way the truncated version of the classical Riesz potentials.

**Definition 4** (*Truncated Riesz potentials*) Let  $\mu$  be Borel measure with finite total mass on  $\mathbb{R}^n$ ; the (truncated) Riesz potential is defined by



$$\mathbf{I}_\beta^\mu(x, R) := \int_0^R \frac{|\mu|(B(x, \varrho))}{\varrho^{n-\beta}} \frac{d\varrho}{\varrho}, \quad \beta > 0,$$

whenever  $x \in \mathbb{R}^n$  and  $0 < R \leq \infty$ .

The word truncated refers to the inequality  $\mathbf{I}_\beta^\mu(x, R) \leq c(n)I_\beta(|\mu|)(x)$ . Now, let us go back to estimates in (5), and let us notice that these cannot hold when  $p \neq 2$ . Indeed, they clearly do not respect the homogeneity properties of the equation. To see this, we consider a non-null solution to  $\operatorname{div}(|Du|^{p-2}Du) = \mu$  with  $p > 2$ , then, for  $\gamma > 0$  we have

$$\tilde{u} := \gamma^{1/(p-1)}u \implies \operatorname{div}(|D\tilde{u}|^{p-2}D\tilde{u}) = \gamma\mu := \tilde{\mu}.$$

Assuming now that the first estimate in (5) would hold, this would yield

$$|\tilde{u}(x)| \lesssim I_2(|\tilde{\mu}|)(x),$$

that is, recalling the definitions of  $\tilde{u}$ ,  $\tilde{\mu}$  and  $I_2$ ,

$$|u(x)| \lesssim \gamma^{(p-2)/(p-1)}I_2(\mu)(x).$$

Upon letting  $\gamma \rightarrow 0$  and using that  $p > 2$ , we would then conclude with  $|u(x)| = 0$  whenever  $I_2(\mu)(x)$  is finite. A similar argument—letting  $\gamma \rightarrow \infty$ —also gives that estimate as in (5) cannot hold in the case  $p < 2$ . In order to overcome this point one is led to consider a new family of potentials. They are obtained by, in some sense, incorporating in Riesz potentials the scaling properties of the  $p$ -Laplacean equation.

**Definition 5** (*Wolff potentials*) Let  $\mu$  be Borel measure with finite total mass on  $\mathbb{R}^n$ ; the nonlinear Wolff potential is defined by

$$\mathbf{W}_{\beta,p}^\mu(x, R) := \int_0^R \left( \frac{|\mu|(B(x, \varrho))}{\varrho^{n-\beta p}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho}, \quad \beta > 0$$

whenever  $x \in \mathbb{R}^n$  and  $0 < R \leq \infty$ .

Wolff potentials, that despite their name were first considered and studied in [28], reduce to Riesz potentials when  $p = 2$ , i.e.  $\mathbf{I}_\beta^\mu \equiv \mathbf{W}_{\beta/2,2}^\mu$ . They play a crucial role in nonlinear potential theory and in the description of the fine properties of solutions to nonlinear equations in divergence form [2, 3, 28, 29, 32, 59, 60].

An important fact about Wolff potentials is that their behaviour can be in several aspects recovered from that of Riesz potentials via so called Havin–Maz’ya potentials  $\mathbf{V}_{\beta,p}(|\mu|)(x)$ . Indeed, following inequality holds provided  $p\beta < n$ :

$$\mathbf{W}_{\beta,p}^\mu(x, \infty) \lesssim I_\beta \left\{ [I_\beta(|\mu|)]^{1/(p-1)} \right\} (x) =: \mathbf{V}_{\beta,p}(|\mu|)(x).$$

The above estimate allows to derive, in a sharp way, almost all types of local estimates starting by the properties of the Riesz potentials, whose action in several function spaces is in fact known; for this we refer to [28, 29].

As first shown in the fundamental works of Kilpeläinen and Malý [37, 38] for the case of nonnegative measures, a neat analog of the first estimate in (5) holds using Wolff potentials. Later on, a new and interesting proof has been offered Trudinger and Wang [65, 66], and this allows to cover the case of general subelliptic operators. Yet different proofs can be found in [39], and in [21], where an approach covering the case of general signed measures has been developed. The final outcome is summarized in the following:

**Theorem 1** [21, 37, 38, 65] *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the equation with measurable coefficients (17) under the assumptions (18). There exists a constant  $c \equiv c(n, p, v, L)$  such that the pointwise estimate*

$$|u(x)| \leq c \mathbf{W}_{1,p}^\mu(x, R) + c \int_{B(x,R)} (|u| + Rs) d\tilde{x} \quad (20)$$

*holds whenever  $B(x, R) \subset \Omega$  and the right hand side is finite. Moreover, if*

$$\lim_{R \rightarrow 0} \mathbf{W}_{1,p}^\mu(x, R) = 0 \quad \text{locally uniformly in } \Omega \text{ w.r.t. } x \quad (21)$$

*then  $u$  is continuous in  $\Omega$ .*

A sketchy proof of this theorem is proposed in Sect. 18 below; extensions to more general operators have been given in [50, 53]. If  $p = 2$ , then  $\mathbf{W}_{1,p}^\mu \equiv \mathbf{I}_2^\mu$  and we retrieve a local analog of the first estimate in (5). A remarkable point here is that estimate (20) is sharp, and the nonlinear potential  $\mathbf{W}_{1,p}^\mu$  cannot be replaced by any other smaller potential. This is in fact reported in the following:

**Theorem 2** [37, 38] *Let  $u \in W^{1,p}(\Omega)$  be a nonnegative weak solution to the Eq. (13) under the assumptions (18), where  $\mu$  is a positive measure and  $s = 0$ . There exists a constant  $c \equiv c(n, p, v, L)$  such that the following pointwise estimate holds whenever  $B(x, 2R) \subset \Omega$  and the Wolff potential is finite:*

$$c^{-1} \mathbf{W}_{1,p}^\mu(x, R) \leq u(x) \leq c \mathbf{W}_{1,p}^\mu(x, 2R) + c \inf_{B(x,R)} u.$$

The possibility of extending pointwise potential estimates to the gradient of solutions has remained an open and discussed issue since the paper [38]. The answer came only recently and here we consider the case of equations of the type (13). The first result in this direction is contained in [57] for the case  $p = 2$ , that is when assumptions (14) are in force. In [57] it is indeed proved the following analog of the second estimate in (5):

$$|D_\xi u(x)| \leq c \mathbf{I}_1^\mu(x, R) + c \int_{B(x,R)} |D_\xi u| d\tilde{x}, \quad \xi \in \{1, \dots, n\}. \quad (22)$$

For the case  $p > 2$  a first result has been given in [21], where the following, apparently natural estimate, has been proved to hold at every Lebesgue point  $x$  of  $Du$ , under assumptions (16):

$$|Du(x)| \leq c \mathbf{W}_{1/p,p}^\mu(x, R) + c \int_{B(x,R)} (|Du| + s) d\tilde{x}. \quad (23)$$

The previous estimate seems to put a final word on the problem of a sharp analog of the second estimate in (5) since the orthodoxy of nonlinear potential theory prescribes that Wolff potentials replace Riesz potentials everywhere when  $p \neq 2$ . Surprisingly enough, in [41] we have shown that this is not the case. Using Wolff potentials is necessary only when estimating solutions, while it is not when passing to their gradients. In fact the following holds:

**Theorem 3** [41] *Let  $u \in W^{1,p}(\Omega)$  be a solution to the Eq. (13), under the assumptions (16). There exists a constant  $c \equiv c(n, p, v, L)$  such that the Riesz potential estimate*

$$|Du(x)|^{p-1} \leq c \mathbf{I}_1^\mu(x, R) + c \left( \int_{B(x,R)} (|Du| + s) d\tilde{x} \right)^{p-1} \quad (24)$$

*holds whenever  $B(x, R) \subset \Omega$  and the right hand side is finite. Moreover, if*

$$\lim_{R \rightarrow 0} \mathbf{I}_1^\mu(x, R) = 0 \text{ locally uniformly in } \Omega \text{ w.r.t. } x \quad (25)$$

*then  $Du$  is continuous in  $\Omega$ .*

The proof of Theorem 3 will be presented in Sect. 15 below, where it will be obtained as a corollary of more general potential estimates. An extension of the previous result to a class of general operators including the  $p$ -Laplacean has been recently given by Baroni in [5].

**Remark 1** An obvious ambiguity arises in the statements of Theorems 1–3 when saying that the related estimates hold whenever the right hand side is finite. This might appear obvious. Another point is that estimates in Theorems 1–3 are stated for every  $x$ , while both  $u(x)$  and  $Du(x)$  are only defined if  $x$  is Lebesgue point, where the so called precise representative can be defined. Both ambiguities are clarified in Sect. 8. There we prove that, both for  $u$  and  $Du$ , the set of Lebesgue points coincides with the one for which inequalities (20) and (24) feature a finite right hand side, respectively.

Observe that estimate (24) obviously improves the one in (23) as, if  $p \geq 2$  then  $1/(p-1) \leq 1$ , and therefore

$$\begin{aligned} \mathbf{I}_1^\mu(x, R) &= \int_0^R \frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \frac{d\varrho}{\varrho} \\ &\lesssim \sum_{i=0}^{\infty} \frac{|\mu|(B(x, R/2^i))}{(R/2^i)^{n-1}} \end{aligned}$$

$$\begin{aligned}
& \lesssim \left[ \sum_{i=0}^{\infty} \left( \frac{|\mu|(B(x, R/2^i))}{(R/2^i)^{n-1}} \right)^{1/(p-1)} \right]^{p-1} \\
& \lesssim \left[ \int_0^{2R} \left( \frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \right)^{1/(p-1)} \frac{d\varrho}{\varrho} \right]^{p-1} \\
& = \left[ \mathbf{W}_{1/p,p}^{\mu}(x, 2R) \right]^{p-1}.
\end{aligned} \tag{26}$$

The implications of Theorem 3 are rather surprising: when switching to the gradient, estimates *linearize* and they reduce to the ones already available for the Poisson equation. The nonlinear, possibly degenerate character of the equations considered plays no role here. In particular, the regularity theory of equations as (13) reduces to that of the Poisson equation up to the  $C^1$ -level. Several facts typical of the linear theory can now be reproduced verbatim. For instance, for solutions  $u \in W^{1,p}(\mathbb{R}^n)$  to (13) defined in the whole  $\mathbb{R}^n$ , the classical estimate via classical Riesz potentials

$$|Du(x)|^{p-1} \leq c \int_{\mathbb{R}^n} \frac{d|\mu|(y)}{|x-y|^{n-1}}$$

follows immediately. For a relevant connection between Theorem 3 and fundamental solutions to general nonlinear measure data problems we refer to Sect. 7.1 below. Another very precise analogy with the classical linear theory is a striking nonlinear extension of a fundamental theorem of Stein [63]. This claims that if  $v \in W^{1,1}$  is a Sobolev function defined in  $\mathbb{R}^n$  with  $n \geq 2$ , then

$$Dv \in L(n, 1) \implies v \text{ is continuous.} \tag{27}$$

This can be regarded as the limiting case of Sobolev–Morrey embedding theorem that reads  $Dv \in L^{n+\varepsilon} \implies v \in C^{0,\varepsilon/(n+\varepsilon)}$  whenever  $\varepsilon > 0$ . Note indeed that  $L^{n+\varepsilon} \subset L(n, 1) \subset L^n$  for every  $\varepsilon > 0$  with all the inclusions being strict. Another way to state Stein’s theorem concerns the regularity of solutions  $u: \Omega \rightarrow \mathbb{R}^N$  to the Laplacean system, and reads

$$\operatorname{div} Du = \Delta u \in L(n, 1) \implies Du \text{ is continuous.}$$

This follows by (27) and classical Calderón–Zygmund theory. The point is now that if  $\mu \in L(n, 1)$ , then a computation involving the definition of norm in Lorentz spaces (see [23, 41] for discussion) implies that (25) takes place, and therefore we conclude with the following:

**Theorem 4** (Nonlinear Stein theorem [41]) *Let  $u \in W^{1,p}(\Omega)$  be a solution to the Eq. (13), under the assumptions (16) and such that  $\mu \in L(n, 1)$  locally in  $\Omega$ . Then  $Du$  is continuous in  $\Omega$ .*

See [12] for a global Lipschitz bound. Without appealing to potentials, but by using different means, the result of the previous theorem also holds for systems.

**Theorem 5** (Vectorial nonlinear Stein theorem [45]) *Let  $u \in W^{1,p}(\Omega, \mathbb{R}^N)$ ,  $N \geq 1$ , be a vector valued solution to the  $p$ -Laplacean system*

$$-\Delta_p u = F.$$

*Assume that the components of the vector field  $F: \Omega \rightarrow \mathbb{R}^N$  locally belong to the space  $L(n, 1)$ . Then  $Du$  is continuous in  $\Omega$ .*

Finally, let us mention that starting from the techniques developed for the last two theorems, similar results can be proved in the case of fully nonlinear equations [15].

## 4 Universal potential estimates

The results of this section have a double aim. On the one hand, they show that estimates (9)–(10) have analogs for nonlinear equations. On the other hand, they show that those in (20), (23) and (24) are actually special cases of a more general class of nonlinear potentials estimates. These, in turn, allow to recover in an optimal way all the basic regularity properties of solutions in terms of the regularity of the assigned datum  $\mu$ . The range of the results implied by such estimates is of course limited by the regularity theory of homogeneous equations as

$$\operatorname{div} a(x, Dw) = 0. \quad (28)$$

Therefore we first recall what is the maximal regularity of solutions to equations as in (28) and in (29) below. In turn, this dramatically changes according to the smoothness assumed on the partial map  $x \rightarrow a(x, \cdot)$ . When measurability is considered, very low degree of regularity is expected and solutions are just Hölder continuous for some exponent. When instead dependence on  $x$  becomes more regular then higher regularity follows. In this case, again for sake of simplicity, and since this does not affect the expositions of the main ideas, we shall confine ourselves to the case of equations with no coefficients as

$$\operatorname{div} a(Dw) = 0. \quad (29)$$

The first result we present upgrades estimate (20) to low order fractional derivatives, and allows to give a sharp formulation of the classical De Giorgi's theory via nonlinear Wolff potential estimates. De Giorgi's theory for equations with measurable provides the existence of a *universal Hölder continuity exponent*  $\alpha_m \in (0, 1)$ , depending only on  $n, p, v, L$ , such that

$$w \in C_{\operatorname{loc}}^{0,\alpha}(\Omega) \quad \text{for every } \alpha < \alpha_m \quad (30)$$

and

$$|w(x) - w(y)| \leq c \int_{B_R} (|w| + Rs) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha. \quad (31)$$

The previous estimate holds whenever  $x, y \in B_{R/2}$  and  $B_R \subset \Omega$ , for a constant depending only on  $n, p, v, L$  and  $\alpha$ . The exponent  $\alpha_m$  can be thought as the *maximal Hölder regularity exponent* associated to the vector field  $a(\cdot)$ . It is universal in the sense that it is independent of  $a(\cdot)$  and of the particular solution considered, but just depends only on  $n, p, v, L$ . For more precise information on  $\alpha_m$  see Theorem 18 and Remark 6 below. It then holds the following:

**Theorem 6** (De Giorgi's theory via potentials [40]) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the equation with measurable coefficients (17) under assumptions (18). Let  $B_R \subset \Omega$  be such that  $x, y \in B_{R/2}$ ; then*

$$|u(x) - u(y)| \leq c \left[ \mathbf{W}_{1-\alpha(p-1)/p,p}^\mu(x, R) + \mathbf{W}_{1-\alpha(p-1)/p,p}^\mu(y, R) \right] |x - y|^\alpha + c \int_{B_R} (|u| + Rs) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha \quad (32)$$

holds provided the right hand side is finite and  $0 \leq \alpha < \alpha_m$ , where the exponent  $\alpha_m$  has been defined in (30)–(31). Moreover, whenever  $\tilde{\alpha} \in [0, \alpha_m)$  is fixed, the dependence of the constant  $c$  is uniform for  $\alpha \in [0, \tilde{\alpha}]$ , in the sense that  $c$  depends only on  $n, p, v, L$  and  $\tilde{\alpha}$ .

Notice that when  $\mu \equiv 0$  estimates (31) and (33) coincide. In general, counterexamples show that in (30)  $\alpha_m \rightarrow 0$  when  $L/v \rightarrow \infty$ . This prevents estimate (33) to hold in general for the full range  $\alpha \in [0, 1)$  when  $x \rightarrow a(x, \cdot)$  is measurable. Things change when the dependence on  $x$  is more regular. For instance, when considering Eq. (29), estimate (31) holds for  $\alpha = 1$ . In this case the validity of (32) extends to the whole interval  $\alpha \in [0, 1]$ . Moreover, estimates (20) and (23) can be uniformly obtained as limiting endpoint cases of (32) for  $\alpha = 0$  and  $\alpha = 1$ , respectively. In fact, it holds the following:

**Theorem 7** (Uniform Wolff potential estimate [40]) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the Eq. (13) under assumptions (16). Let  $B_R \subset \Omega$  be such that  $x, y \in B_{R/2}$ ; then estimate (32) holds for  $\alpha \in [0, 1]$ , whenever the right hand side is finite. The constant  $c$  depends only on  $n, p, v, L$  but is otherwise independent of  $\alpha$ .*

Estimate (32) does not catch up, when  $\alpha \rightarrow 1$ , the optimal gradient bound in (24). There is a natural reason for such a lack of endpoint property: it catches the other optimal bound (20) as  $\alpha \rightarrow 0$ , and the two cases involve different potentials. In fact, the change in the nature of the estimates when passing from (20) to (24) requires another theorem, parallel to Theorem 7.

**Theorem 8** (Uniform Riesz potential estimate) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the Eq. (13) under assumptions (16). Let  $B_R \subset \Omega$  be such that  $x, y \in B_{R/4}$ ; then*

$$|u(x) - u(y)| \leq c \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) + \mathbf{I}_{p-\alpha(p-1)}^\mu(y, R) \right]^{1/(p-1)} |x - y|^\alpha + c \int_{B_R} (|u| + Rs) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha \quad (33)$$

holds provided the right hand side is finite and  $0 < \alpha \leq 1$ . Moreover, whenever  $\tilde{\alpha} \in (0, 1]$  is fixed, the dependence of the constant  $c$  is uniform for  $\alpha \in [\tilde{\alpha}, 1]$  as  $c$  depends only  $n, p, v, L$  and  $\tilde{\alpha}$ .

The previous result is here presented for the first time and for the proof we refer to Sect. 5 below. There Theorem 8 is actually obtained as a corollary of a more general estimate for certain maximal operators of  $Du$ . From the discussion made there it will be clear that the points  $x, y$  are Lebesgue points of  $u$  when the right hand side in (33) is finite; compare with Remark 1. Theorems 6 and 8 provide altogether the optimal analog of estimate (9) and in fact they both unify with Theorem 7 when  $p = 2$ , when Wolff and Riesz potentials do coincide. We remark that is also possible to quantify the blow-up of the constant  $c$  in (33) as  $\alpha \rightarrow 0$ ; see Remark 7 below.

We now examine the situation for the gradient, giving an estimate which is again bound to give back (10) when  $p = 2$ . As already done before Theorem 6, we recall the basic information about the maximal regularity of solutions to homogeneous equations as in (29). These in turn involve the fundamentals of the regularity theory of the  $p$ -Laplacean equations and can be summarised by saying that there exists a positive exponent  $\alpha_M \in (0, 1)$ , depending only on  $n, p, v$  and  $L$ , such that whenever  $x, y \in B_{R/2}$ ,

$$Dw \in C_{\text{loc}}^{0,\alpha}(\Omega) \quad \text{for every } \alpha < \alpha_M \quad (34)$$

and

$$|Dw(x) - Dw(y)| \leq c \int_{B_R} (|Dw| + s) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha \quad (35)$$

hold for any local solution  $w$  to (29). The constant  $c$  depends only on  $n, p, v, L$  and  $\alpha$ . For estimate (35) and the exponent  $\alpha_M$  see Theorem 17 and Remark 5 below. The result in (34) finds its origins in the basic work of Uraltseva [67] while further approaches can be found in [18, 24, 46, 51]. We remark that all these results are again based on the groundbreaking work of De Giorgi [16]. The exponent  $\alpha_M$  is in general strictly less than one (see again [67]); while lower bounds on  $\alpha_M$  can be obtained by tracking the constant dependence in the various proofs, its precise (optimal) value is not known. An optimal lower bound for  $\alpha_M$  is conjectured to be  $1/3$ , according to the regularity exhibited by the solutions to the so called  $\infty$ -Laplacean equation [4]. By taking (34) into account we have the following nonlinear version of estimate (10):

**Theorem 9** (Uniform Wolff potential estimate) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the Eq. (13) under assumptions (16). Let  $B_R \subset \Omega$  be such that  $x, y \in B_{R/4}$ ; then*

$$\begin{aligned} |Du(x) - Du(y)| &\leq c \left[ \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p},p}^\mu(x, R) + \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p},p}^\mu(y, R) \right] |x - y|^\alpha \\ &\quad + c \int_{B_R} |Du - (Du)_{B_R}| d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha \end{aligned} \quad (36)$$

holds provided the right hand side is finite and

$$0 \leq \alpha < \min\{1/(p-1), \alpha_M\}. \quad (37)$$

The exponent  $\alpha_M$  has been defined in (34). Moreover, whenever  $0 \leq \tilde{\alpha} < \min\{1/(p-1), \alpha_M\}$  is fixed, the dependence of the constant  $c$  is uniform for  $\alpha \in [0, \tilde{\alpha}]$ , in the sense that  $c$  depends only  $n, p, v, L$  and  $\tilde{\alpha}$ .

The proof of this theorem will be presented in Sect. 16 below. Notice that the limitation in (37) makes the statement of Theorem 9 consistent with the definition of Wolff potential given in Definition 5, where it must be  $\beta > 0$ .

## 5 Maximal-potential estimates

Let us now explain Theorem 8 proceeding for some while with a few purely heuristic arguments. By considering the function  $m$  appearing in (11) as a fractional derivative of order  $\alpha$ , we can express the content of Theorem 8 in the following way:

$$|\partial^\alpha u(x)|^{p-1} \lesssim I_{p-\alpha(p-1)}(|\mu|)(x), \quad 0 < \alpha \leq 1 \quad (38)$$

thereby obtaining a nonlinear version of (12). The singularity of the case  $\alpha = 0$  essentially stems from the fact that Wolff potentials come then into the play when considering pointwise estimates for  $u$ . Let us proceed heuristically and see why Riesz potentials intervene as long as fractional derivatives are considered. We rewrite Eq. (15) as the decoupled system

$$\begin{cases} -\operatorname{div} H = \mu \\ H = |Du|^{p-2} Du. \end{cases}$$

This viewpoint tells us that an equation as in (15), which is commonly seen as a *nonlinear equation in the gradient*, can be also seen as linear equation in a *nonlinear vector field of the gradient*, that is  $H$ . The analysis of equations as  $-\operatorname{div} H = \mu$ , being  $f \rightarrow \operatorname{div} f$  a differential operator of order one, typically involves estimates of  $H$  via operators as  $\mathbf{I}_1^\mu$ . In turn this explains the appearance of estimates as in (24). The reader will immediately recognize that no similar argument applies to  $u$ , as the equation depends directly on  $Du$  and not  $u$ . This is ultimately the reason for the appearance of Wolff potentials when deriving pointwise for  $u$  as (20). Let us now go to intermediate derivatives and let us believe that fractional derivatives  $\partial^\alpha u$  can be controlled by in turn controlling quantities as

$$\frac{|u - (u)_{B_R}|}{R^\alpha} \approx R^{1-\alpha} |Du| \approx |\partial^\alpha u| \quad (39)$$

on every possible ball  $B_R$ . Using a typical dimension analysis viewpoint, the appearance in the estimates of a multiplicative factor as  $R^{-\beta}$  amounts to consider a corresponding derivative of order  $\beta$  in the estimates. Now, let us rewrite Eq. (15) as

$$R^{(\alpha-1)(p-1)} \operatorname{div} (|R^{1-\alpha} Du|^{p-2} R^{1-\alpha} Du) = \mu$$

and note that in this scheme of reasoning the operator  $f \rightarrow R^{(\alpha-1)(p-1)} \operatorname{div} f$  has now order  $p - \alpha(p-1)$ . Its inversion should therefore involve potentials as  $\mathbf{I}_{p-\alpha(p-1)}^\mu$ . At



this point, recalling the identification (39) leads to conclude with an estimate with the dimensional meaning of (38). The point is now to make everything rigorous. A natural way to make the identification in (39) is to use a nonlocal operator as for instance a fractional maximal operator. It is therefore time for the following:

**Definition 6** (*Fractional maximal operator*) Let  $\beta \in [0, n]$ ,  $x \in \Omega$  and  $R < \text{dist}(x, \partial\Omega)$ , and let  $f$  be an  $L^1(\Omega)$ -function or a measure with finite mass; the function defined by

$$M_{\beta,R}(f)(x) := \sup_{0 < r \leq R} r^\beta \frac{|f|(B(x, r))}{|B(x, r)|}$$

is called the restricted (centered) fractional  $\beta$  maximal function of  $f$ .

Using fractional maximal operators allows to get Riesz potential estimates which are uniform in the whole range  $[0, 1]$ , and that provide a rigorous interpretation of previous heuristic arguments. The outcome is summarized in the following theorem, which appears here for the first time. The proof is contained in Sect. 12 below (see Sect. 14.4 for the case of SOLA).

**Theorem 10** (Uniform maximal-potential estimates) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the Eq. (13) under assumptions (16). Then for every ball  $B(x, R) \subset \Omega$  the following estimate:*

$$M_{1-\alpha,R}(Du)(x) \leq c \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) \right]^{1/(p-1)} + cR^{1-\alpha} \int_{B_R} (|Du| + s) d\tilde{x} \quad (40)$$

holds uniformly in  $\alpha \in [0, 1]$ , with  $c \equiv c(n, p, \nu, L)$ .

The maximal operators of Definition 3 naturally connect to those presented in Definition 6 via Poincaré inequality

$$M_{\alpha,R}^\#(u)(x) \leq cM_{1-\alpha,R}(Du)(x), \quad \alpha \in [0, 1] \quad (41)$$

so that a uniform estimate on  $M_{\alpha,R}^\#(u)$  follows appealing to Theorem 10, that is

$$M_{\alpha,R}^\#(u)(x) \leq c \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) \right]^{1/(p-1)} + cR^{1-\alpha} \int_{B_R} (|Du| + s) d\tilde{x}. \quad (42)$$

We yet remark that in the estimate in the last display the constant  $c$  is independent of  $\alpha$ , and in fact the estimate holds uniformly for  $\alpha \in [0, 1]$ .

The turning point to pointwise estimates is now given by the fact that the sharp maximal operator  $M_{\alpha,R}^\#(u)(x)$  controls the pointwise behaviour of  $u$  provided  $\alpha > 0$ . This fact is expressed in the following Proposition, a first form of which is present in [17]. It in turn relies on some original arguments of Campanato [10].

**Proposition 1** *Let  $f \in L^1(B_{8R/5})$ ; for every  $\alpha \in (0, 1]$  the inequality*

$$|f(x) - f(y)| \leq \frac{c}{\alpha} \left[ M_{\alpha,R}^\#(f)(x) + M_{\alpha,R}^\#(f)(y) \right] |x - y|^\alpha \quad (43)$$

holds whenever  $x, y \in B_{2R/5}$ , for a constant  $c$  depending only on  $n$ . More precisely,  $x$  and  $y$  are Lebesgue points of  $f$  whenever  $M_{\alpha,R}^\#(f)(x)$  and  $M_{\alpha,R}^\#(f)(y)$  are finite, respectively. Therefore, whenever the right hand side in (43) is finite, the values of  $f$  are defined as follows:

$$f(x) := \lim_{\varrho \rightarrow 0} (f)_{B(x,\varrho)} \quad \text{and} \quad f(y) := \lim_{\varrho \rightarrow 0} (f)_{B(y,\varrho)}.$$

Proposition 1 allows to obtain Theorem 8 as a corollary of Theorem 10 and in particular of (42); see Sect. 12.7 below for the proofs, including the one of Proposition 1. We believe that Theorem 10 helps to understand the peculiar nature of the case  $\alpha = 0$  and the occurrence of Wolff potentials instead of Riesz potentials. Indeed, an  $L^\infty$ -bound on the sharp maximal operator is essentially equivalent to require that  $u$  belongs to the space BMO (introduced in [34], see (68) below). On the other hand, a real  $L^\infty$ -bound on  $u$  itself requires an  $L^\infty$ -bound on  $\mathbf{W}_{1,p}^\mu$ . This situation looks completely natural when considering typical examples of functions lying in  $\text{BMO} \setminus L^\infty$ , as for instance  $\log |x|$ . The log-scale is indeed the one making a difference between a potentials as  $\mathbf{W}_{1,p}^\mu$  and  $\mathbf{I}_1^\mu$ . Taking for instance

$$\mu(x) = \frac{1}{|x|^p \log^\beta |x|}$$

we see that  $\mathbf{I}_p^\mu(0, 1) < \infty$  provided  $\beta > 1$  while  $\mathbf{W}_{1,p}^\mu(0, 1) < \infty$  provided  $\beta > p - 1$ . In the next section we shall in fact describe a few additional results showing that also the case  $\alpha = 1$  reveals to be special when using maximal operators instead of potentials.

### 5.1 Non-endpoint maximal estimates

An interesting fact concerning the two cases  $\alpha = 0, 1$  of (40) is that when looking for bounds on intermediate derivatives, potentials can be replaced by maximal operators. These give in fact smaller quantities. Maximal operators, as we shall see, allow to replace conditions typically given in Lebesgue or Lorentz spaces, with weaker conditions formulated in terms of Marcinkiewicz spaces. Let us recall the following, elementary inequality which holds for  $\gamma \in (0, 1)$

$$M_{\beta,\gamma R}(\mu)(x) \leq c(n, \gamma, \beta) \mathbf{I}_\beta^\mu(x, R),$$

see for instance [40, Lemma 4.1]. Then we have the following:

**Theorem 11** (Intermediate maximal estimates [40]) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to (13) under the assumptions (16). Let  $B_R \subset \Omega$  be a ball centred at  $x$ ; then the estimate*

$$\begin{aligned} & M_{\alpha,R}^\#(u)(x) + M_{1-\alpha,R}(Du)(x) \\ & \leq c \left[ M_{p-\alpha(p-1),R}(\mu)(x) \right]^{1/(p-1)} + cR^{-\alpha} \int_{B_R} |u - (u)_{B_R}| d\tilde{x} + cR^{1-\alpha} s \end{aligned} \quad (44)$$

holds uniformly in  $\alpha \in [0, \tilde{\alpha}]$ , whenever  $0 < \tilde{\alpha} < 1$ . The constant  $c$  depends only on  $n, p, v, L, \tilde{\alpha}$ .

As for gradient oscillations, we instead have the following result, whose proof is contained in Sect. 13 below.

**Theorem 12** (Gradient sharp maximal estimate [40]) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to (13) under the assumptions (16). Let  $B_R \subset \Omega$  be a ball centered at  $x$ ; then the following estimate:*

$$M_{\alpha,R}^{\#}(Du)(x) \leq c \left[ M_{1-\alpha(p-1),R}(\mu)(x) \right]^{1/(p-1)} + cR^{-\alpha} \int_{B_R} |Du - (Du)_{B_R}| d\tilde{x}$$

holds uniformly in  $\alpha \in [0, \tilde{\alpha}]$  whenever

$$0 \leq \tilde{\alpha} < \min\{1/(p-1), \alpha_M\},$$

for a constant  $c$  depending only on  $n, p, v, L, \tilde{\alpha}$ . The exponent  $\alpha_M$  has been introduced in (34)–(35).

## 6 Equations with coefficients

Getting nonlinear potential estimates catching regularity beyond Theorem 6 for equations as in (17), necessitates to assume more regularity on the partial map  $x \mapsto a(x, \cdot)$ . Keep in mind the discussion before Theorem 6. We therefore define the following averaged vector field:

$$(a)_{x,r}(z) = \int_{B(x,r)} a(\tilde{x}, z) d\tilde{x}, \quad \text{for } z \in \mathbb{R}^n, \quad (45)$$

whenever  $B(x, r) \subset \Omega$ , and then the averaged (and renormalized) modulus of continuity of  $x \mapsto a(x, \cdot)$  as follows:

$$\omega(r) := \left[ \sup_{z \in \mathbb{R}^n, B(x,r) \subset \Omega} \int_{B(x,r)} \left( \frac{|a(\tilde{x}, z) - (a)_{x,r}(z)|}{(|z| + s)^{p-1}} \right)^2 d\tilde{x} \right]^{1/2}. \quad (46)$$

When considering equations of the type

$$-\operatorname{div}(c(x)\tilde{a}(Du)) = \mu \quad (47)$$

the definition in (46) gives something which is comparable to the usual modulus of continuity of the function  $c(\cdot)$ . We then have

**Theorem 13** *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to (17), with assumptions (16) (uniformly) verified by the partial map  $z \rightarrow a(\cdot, z)$  for every  $x \in \Omega$ . Then*

- [40] For every  $\tilde{\alpha} \in (0, 1)$  there exists a number  $\delta \in (0, 1)$ , depending only on  $n, p, v, L, \tilde{\alpha}$ , such that

$$\limsup_{\varrho \rightarrow 0} \omega(\varrho) \leq \delta \quad (48)$$

implies the validity of estimate (32) whenever  $\alpha \in [0, \alpha_m)$ . In particular, if the limit in (48) is zero estimate (32) holds for every  $\alpha < 1$

- [6] If the partial map  $x \rightarrow a(x, \cdot)$  is Dini-continuous, i.e.

$$\int_0^r \omega(\varrho) \frac{d\varrho}{\varrho} < \infty, \quad (49)$$

then estimates (33) and (40) continue to hold, for  $\alpha \in [\tilde{\alpha}, 1]$  and  $\alpha \in [0, 1]$ , respectively

- [40] If the partial map  $x \rightarrow a(x, \cdot)$  is Hölder continuous with exponent  $\tilde{\alpha} < \min\{1/(p-1), \alpha_M\}$  in the sense that

$$\sup_r \int_0^r \frac{[\omega(\varrho)]^{2/p}}{\varrho^{\tilde{\alpha}}} \frac{d\varrho}{\varrho} < \infty, \quad (50)$$

then estimate (36) holds whenever  $\alpha < \tilde{\alpha}$ .

In all three cases the constants involved in the estimates additionally depend on the quantities appearing in (48)–(50).

Let us now briefly comment on the assumptions made in the last theorem, remarking that all of them are essentially necessary. Assumption (48) allows to catch the case in which, when referring to Eq. (47), the function  $c(\cdot)$  is not even continuous, but of class VMO or BMO with small seminorm. This reconnects the theory presented here to the classical Calderón–Zygmund theory, where VMO-regularity of coefficients is required to establish integrability estimates related to estimate (32); see for instance [9]. Assumption (149) is necessary too. Indeed, estimate (33) implies gradient boundedness when  $\mu$  is good enough, while Dini continuity of coefficients is known to be necessary to get Lipschitz regularity of solutions; see [33]. The proof of the second statement in the last theorem has the techniques developed for Theorem 10 as starting point, and will appear in the forthcoming paper [6]. Finally, assumption (50) seems to be natural in view of the usual Schauder theory. This prescribes that, in order to have Hölder continuity of the gradient, the Hölder continuity of  $x \mapsto a(x, \cdot)$  is necessary.

## 7 Interlude on measure data problems

When looking at equations as in (17) we have considered distributional solutions lying in  $W^{1,p}$ . This choice is actually aimed at simplifying the presentation in a way that emphasises the results obtained in the form of a priori estimates. Dealing with  $W^{1,p}$ -solutions is anyway very much restrictive in view of the fact that, typically,

distributional solutions to measure data problems do not enjoy this regularity. A typical instance is the so called *nonlinear Green's function*

$$G_p(x) \approx \begin{cases} \left(|x|^{\frac{p-n}{p-1}} - 1\right) & \text{if } 1 < p \neq n \\ \log |x| & \text{if } p = n \end{cases} \quad (51)$$

which solves, in the sense of distributions, the problem

$$\begin{cases} -\Delta_p u = \delta & \text{in } B(0, 1) \\ u = 0 & \text{on } \partial B(0, 1). \end{cases} \quad (52)$$

In the right hand side of (52)<sub>1</sub> there appears the Dirac measure  $\delta$  charging the origin. A straightforward calculation reveals that  $G_p \notin W^{1,p}$  while it holds that

$$|DG_p|^{p-1} \in \mathcal{M}^{n/(n-1)} \iff |\{|DG_p|^{p-1} > \lambda\}| \lesssim \lambda^{-n/(n-1)}. \quad (53)$$

We recall here that the Marcinkiewicz space  $\mathcal{M}^q(\Omega)$ , for  $1 \leq q < \infty$ , is the set of all measurable maps  $\mu$  satisfying the condition

$$\sup_{\lambda \geq 0} \lambda^q |\{x : |\mu|(x) > \lambda\}| < \infty. \quad (54)$$

The above discussion now motivates the following:

**Definition 7** (*Very weak solutions*) A function  $u \in W_{\text{loc}}^{1,1}(\Omega)$  is called a very weak solution  $u$  to the Eq. (17) in  $\Omega$  if  $a(x, Du) \in L_{\text{loc}}^1(\Omega, \mathbb{R}^n)$  and (19) holds for every  $\varphi \in C_c^\infty(\Omega)$ .

The terminology “very weak” is used to emphasize the fact that such solutions do not belong, in general, to the natural Sobolev space  $W^{1,p}(\Omega)$ , i.e. they are not energy solutions. Note that, when considering assumptions (18), in order to guarantee that  $a(x, Du) \in L^1(\Omega, \mathbb{R}^n)$  it suffices, for instance, to have  $Du \in W^{1,p-1}(\Omega)$ . Definition 7, not surprisingly, poses problems. For instance, very weak solutions may exist beside usual energy solutions [61], and are therefore not in general unique. As a matter of fact one of the basic open issues of the theory of measure data problems is to find a function class where to solve in a unique way Dirichlet problems of the type

$$\begin{cases} -\operatorname{div} a(x, Du) = \mu & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (55)$$

with  $\mu$  being in the most general case a (signed) Borel measure with finite total mass. We will not pursue this matter here, rather referring to [14, 36] for a more comprehensive discussion. Here we are interested in a class of solutions which we regard to be very natural (see in fact the equivalence results obtained in [36]) and for which all the potential estimates described in this paper continue to hold. This is the class of SOLA (Solutions Obtained by Limits of Approximations), which is introduced in the

**Definition 8** (SOLA) A function  $u \in W_{\text{loc}}^{1,1}(\Omega)$  is a SOLA to the Eq. (17) iff is a very weak solution and there exists a sequence of local energy solutions  $\{u_k\} \subset W_{\text{loc}}^{1,p}(\Omega)$  to the equations

$$-\operatorname{div} a(x, Du_k) = \mu_k,$$

such that  $u_k \rightarrow u$  locally in  $W^{1,p-1}(\Omega)$ .

SOLA are special because they are selected via an approximation procedure using more regular energy solutions, and thereby they inherit a few of their basic properties. For instance: their precise representative is defined out of a null  $p$ -capacity set, exactly as  $W^{1,p}$ -functions; see Theorem 16 below. Let us recall how to build a SOLA to (55), following [8]. One considers solutions  $u_k \in W_0^{1,p}(\Omega)$  to regularised problems

$$\begin{cases} -\operatorname{div} a(x, Du_k) = \mu_k & \text{in } \Omega \\ u_k = 0 & \text{on } \partial\Omega. \end{cases}$$

The right hand sides data  $\mu_k := \mu * \phi_k \in C^\infty$  are canonically obtained by smoothing  $\mu$  via convolution with a sequence of smooth standard, smooth mollifiers  $\{\phi_k\}$ . This approach leads to the existence of a very weak solution  $u \in W_0^{1,p-1}(\Omega)$  such that  $u_k \rightarrow u$  in  $W_0^{1,p-1}(\Omega)$  (up to a not relabeled subsequence). SOLA are not known to be unique, except in a few special cases (for instance when  $\mu \in L^1$ ). See [8, 13, 14, 36] for a larger discussion. We summarise the basic existence and regularity results available for SOLA in the following:

**Theorem 14** [8, 55] *Under the assumptions (18) with  $p \leq n$ , there exists a SOLA  $u \in W_0^{1,p-1}(\Omega)$  to (55). Moreover, every SOLA  $u \in W_{\text{loc}}^{1,1}(\Omega)$  to (17) is such that*

$$u \in W_{\text{loc}}^{1,q}(\Omega) \text{ for every } q < n(p-1)/(n-1). \quad (56)$$

*In the case of equations of the type (13) under assumptions (16), we also have*

$$Du \in W_{\text{loc}}^{(1-\varepsilon)/(p-1), p-1}(\Omega) \text{ for every } \varepsilon \in (0, 1). \quad (57)$$

A preliminary integrability result was obtained for a different kind of solutions in the pioneering paper of Lindqvist [48]. The restriction  $p \leq n$  is aimed at focusing on the case of SOLA, since when  $p > n$  then  $\mu$  belongs to the dual of  $W^{1,p}$  and it is possible to consider standard energy solutions. In display (57) a fractional Sobolev space appears. The meaning is actually that

$$\int_A \int_A \frac{|Du(x) - Du(y)|^{p-1}}{|x-y|^{n+1-\varepsilon}} dx dy < \infty \quad \text{for every } \varepsilon \in (0, 1)$$

holds whenever  $A \Subset \Omega$ . We also remark that, in fact, it can be proved that every SOLA  $u$  is such that

$$|Du|^{p-1} \in \mathcal{M}^{n/(n-1)}(\Omega). \quad (58)$$

This means that every SOLA exhibits exactly the same integrability of the nonlinear Green's function  $G_p$  described in (53) (see [7, 20, 55]). Both the results in (56) and (57) are sharp, as follows again considering  $G_p$ , which can be in fact proved to be the only SOLA to the problem in (52). This last fact follows combining the results in [35, 62]. As mentioned above, by mean of approximations arguments, *all the potential estimate in this paper continue to hold for any SOLA to (55)*. More details are in Sect. 14.4 below, where we in particular discuss the possibility of characterizing the Lebesgue points of SOLA via linear and nonlinear potentials.

### 7.1 Potential estimates and fundamental solutions

To check to which extent estimates (20) and (24) replace the usual linear representation formulas in the nonlinear case, the best thing to do is to see how they reproduce the behaviour of the nonlinear fundamental solution  $G_p$ . We already know the optimality of estimate (20) from Theorem 2. We therefore concentrate on (24) and show that it actually reverses when considering the nonlinear Green's function  $G_p$ . This means that for  $x \in B(0, 1) \setminus \{0\}$  we have

$$\int_{B(x, R)} |DG_p| d\tilde{x} + [\mathbf{I}_1^\delta(x, R)]^{1/(p-1)} \leq c |DG_p(x)|, \quad (59)$$

with  $R = 2|x|$ , and for a constant  $c$  depending only on  $n$  and  $p$ . Indeed by coarea formula we have

$$\begin{aligned} \int_{B(x, R)} |DG_p| d\tilde{x} &\leq c \int_{B(0, 2R)} |DG_p| d\tilde{x} \\ &= \frac{c(n, p)}{R^n} \int_0^{2R} \int_{\partial B_\varrho} |DG_p(y)| d\mathcal{H}^{n-1}(y) d\varrho \\ &\leq \frac{c(n, p)}{R^n} \int_0^{2R} \varrho^{n+(1-n)/(p-1)} \frac{d\varrho}{\varrho} \\ &\leq c(n, p) R^{(1-n)/(p-1)} \leq c(n, p) |DG_p(x)|. \end{aligned}$$

On the other hand, as  $\delta(B(x, \varrho)) = 0$  for  $\varrho \leq R/2$  and  $\delta(B(x, \varrho)) = 1$  otherwise, we also have

$$\mathbf{I}_1^\delta(x, R) = \int_{R/2}^R \frac{d\varrho}{\varrho^n} = \frac{2^n - 1}{n - 1} R^{1-n} \leq c(n, p) |DG_p(x)|^{p-1},$$

so that (59) follows combining the last two inequalities.

## 8 Fine properties of solutions via potentials

We have seen that linear and nonlinear potentials locally control the behaviour of solutions. It is at this point not surprising to discover that potentials also control their so called fine properties. In this respect we present two theorems. The former is concerned with the pointwise behaviour of gradients of SOLA, and in particular with their Lebesgue points. This result employs Riesz potentials. The latter is instead concerned with Lebesgue points of solutions and uses Wolff potentials.

**Theorem 15** *Theorem 3 continues to hold whenever  $u \in W^{1,p-1}(\Omega)$  is a SOLA to (13). Moreover, the condition*

$$\mathbf{I}_1^\mu(x, R) < \infty \quad (60)$$

*implies that the following limit exists and therefore defines the precise representative of  $Du$  at the point  $x$ :*

$$\lim_{\varrho \rightarrow 0} (Du)_{B(x, \varrho)} =: Du(x). \quad (61)$$

The proof of this last theorem is included in Sect. 17 below. The analogous statement for solutions is the following:

**Theorem 16** *Theorem 1 continues to hold whenever  $u \in W^{1,p-1}(\Omega)$  is a SOLA to (17). Moreover, the condition*

$$\mathbf{W}_{1,p}^\mu(x, R) < \infty \quad (62)$$

*implies that the following limit exists and therefore defines the precise representative of  $u$  at the point  $x$ :*

$$\lim_{\varrho \rightarrow 0} (u)_{B(x, \varrho)} =: u(x). \quad (63)$$

*In particular, the set of non-Lebesgue points of  $u$  has  $p$ -capacity zero.*

We notice that the assertion about the  $p$ -capacity follows directly from the fact that set of points where the Wolff potential  $\mathbf{W}_{1,p}^\mu$  blows-up has zero  $p$ -capacity (see for instance [30]).

**Remark 2** (Hausdorff dimension of singular sets of SOLA) Theorem 16 allows to define SOLA—which are initially defined only almost everywhere via convergence—outside a singular set (i.e. set of non-Lebesgue points) of Hausdorff dimension not larger than  $n - p$ , when  $p \leq n$ . This establishes a connection with another class of solutions to measure data problem. These are called  $p$ -superharmonic functions and are defined in the case the measure  $\mu$  is nonnegative; see [30, 38]. For such solutions every point is a Lebesgue point, by construction. The connection is now given by the fact that, in view of [36], every SOLA has a superharmonic representative (that is,



they coincide almost everywhere) whenever the measure is nonnegative. In the case of Theorem 15 we instead can conclude that the Hausdorff dimension of the singular set of  $Du$  is not larger than  $n - 1$ . The last estimate follows on the other hand also by (57). Indeed, the Hausdorff dimension of the set of non-Lebesgue points of a general  $W^{s,\gamma}$ -map, with  $s\gamma < n$ , has Hausdorff dimension not larger than  $n - s\gamma$ ; see [54].

*Remark 3* (Analogies with linear potential theory) Theorems 15 and 16 can be also considered as analogs of classical facts in linear potential theory, concerning the pointwise behaviour of solutions to the Poisson equation  $-\Delta u = \mu$ . As usual, in the linear case they follow from explicit representation formulas and abstract analysis of potentials. Here they are replaced by nonlinear potential estimates. Indeed, define

$$f(x) := (G_\alpha * g)(x) := \int_{\mathbb{R}^n} G_\alpha(x - y)g(y) dy,$$

where  $g \in L^q(\mathbb{R}^n)$  with  $q > 1$ . Here  $G_\alpha$  denotes the  $\alpha$ -Bessel potential, that is defined via

$$G_\alpha := \mathcal{F}^{-1} \left( (1 + |\xi|^2)^{-\alpha/2} \right),$$

where  $\mathcal{F}$  denotes the Fourier transform and  $\alpha \in (0, n)$ . Then the limit

$$\lim_{\varrho \rightarrow 0} (f)_{B(x, \varrho)} =: f(x)$$

exists—and thereby defines the precise representative of  $f$ —up to a set of zero  $(\alpha, q)$ -capacity. For these facts we refer to [3], and in particular to [3, Proposition 6.1.3]

## 9 Regularity and corollaries

The estimates presented up to now allow to give a comprehensive and unified picture of the regularity results available for quasilinear equations. We will now briefly describe a few consequences of such estimates. First of all, let us recall that combining Theorem 11 and Proposition 1 yields the following pointwise inequality:

$$\begin{aligned} |u(x) - u(y)| &\leq c \left[ M_{p-\alpha(p-1), R}(\mu)(x) + M_{p-\alpha(p-1), R}(\mu)(y) \right]^{1/(p-1)} |x - y|^\alpha \\ &\quad + c \int_{B_R} (|u| + Rs) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha, \end{aligned} \quad (64)$$

that holds whenever  $x, y \in B_{R/4}$ ,  $B_R \subset \Omega$ ,  $\alpha \in (0, 1)$ , where  $c \equiv c(n, p, v, L, \alpha)$ . Notice here that the blow-up of  $c$  takes place as soon as  $\alpha$  approaches 0 or 1. In a completely similar way, by Theorem 12 and again Proposition 1, we then have that the inequality

$$|Du(x) - Du(y)| \leq c \left[ M_{1-\alpha(p-1),R}(\mu)(x) + M_{1-\alpha(p-1),R}(\mu)(y) \right]^{1/(p-1)} |x - y|^\alpha \\ + c \int_{B_R} (|Du| + s) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^\alpha \quad (65)$$

holds with  $x, y, r, B_R$  and  $c$  as above, but for  $\alpha \in (0, \alpha_M)$ . This time the constant  $c$  blows-up when  $\alpha \rightarrow 0$  or  $\alpha \rightarrow \alpha_M$ . We are now interested in seeing how the results presented up to now, including estimates (64)–(65), imply regularity of solutions in various relevant function spaces. Let us briefly recall that the Morrey space  $L^{q,\theta}(\Omega)$  for  $q \geq 1$  and  $0 \leq \theta \leq n$ , is defined by requiring that its member  $f \in L^q(\Omega)$  satisfies

$$\sup_{B_\varrho \subset \Omega} \varrho^{\theta-n} \int_{B_\varrho} |f|^q d\tilde{x} < \infty, \quad (66)$$

while their local version is defined in the usual way. We refer to [1, 56] for more on Morrey spaces and their basic properties with related references. We now have the following, comprehensive:

**Corollary 1** *Let  $u \in W^{1,p}(\Omega)$  be a weak solution—or a SOLA—to the Eq. (13) under assumptions (16). Then, with all the spaces being meant in their local versions, we have that*

- (C1) *if  $\alpha \in (0, 1)$  and  $q > n(p-1)/[n-p+\alpha(p-1)]$ , then*

$$\mu \in L^{\frac{nq}{n(p-1)+[p-\alpha(p-1)]q}} \implies u \in C_q^\alpha$$

*(estimate (64))*

- (C2) *if  $p < n$ , then*

$$\mu \in \mathcal{M}^{n/p} \implies |\mu|(B_\varrho) \lesssim \varrho^{n-p} \implies u \in BMO$$

*and*

$$p = n \implies u \in BMO$$

*(Theorem 11 with  $\alpha = 0$ )*

- (C3)  $\mathbf{W}_{1,p}^\mu \in L^\infty \implies u \in L^\infty$  *(Theorem 1)*
- (C4) *if  $p < n$ , then  $\mu \in L(n/p, 1/(p-1)) \implies u \in C^0$  (Theorem 1)*
- (C5) *if  $\alpha \in (0, 1)$  and  $p \leq n$ , then*

$$\mu \in \mathcal{M}^{\frac{n}{p-\alpha(p-1)}} \implies |\mu|(B_\varrho) \lesssim \varrho^{n-p+\alpha(p-1)} \implies u \in C^{0,\alpha}$$

*(estimate (64))*

- (C6) *if  $p \leq n$ , then  $\mu \in \mathcal{M}_b \implies Du \in \mathcal{M}^{n(p-1)/(n-1)}$  (Theorem 3)*
- (C7) *if  $1 < q < n$ , then*

$$\mu \in L(q, \gamma) \implies Du \in L\left(\frac{nq(p-1)}{n-q}, \gamma(p-1)\right)$$

*(Theorem 3)*

- (C8) if  $1 < q < \theta$ , then  $|Du|^{p-1} \in L^{\theta q/(\theta-q), \theta}$  (Theorem 3)
- (C9)  $\mu \in \mathcal{M}^n \implies |\mu|(B_\varrho) \lesssim \varrho^{n-1} \implies Du \in \text{BMO}$  (Theorem 12)
- (C10)  $\mathbf{I}_1^\mu \in L^\infty \implies Du \in L^\infty$  (Theorem 8)
- (C11)  $\mu \in L(n, 1) \implies Du \in C^0$  (Theorem 3)
- (C12) if  $0 < \alpha < \min\{1/(p-1), \alpha_M\}$ , then

$$\mu \in \mathcal{M}^{\frac{n}{1-\alpha(p-1)}} \implies |\mu|(B_\varrho) \lesssim \varrho^{n-1+\alpha(p-1)} \implies Du \in C^{0,\alpha}$$

(estimate (65)).

*Proof* In order to prove (C1), we begin recalling the following mapping property of the fractional maximal operator:

$$M_{\beta,R}: L^t \rightarrow L^{nt/(n-\beta t)}, \quad (67)$$

that holds whenever  $1 < t < n/\beta$  (see [26, 56]). Then we go to estimate (64) and yet recall the definition of Calderón spaces given in Definition 2. This leads to determine

$$\frac{nt}{n - [p - \alpha(p-1)]t} = \frac{q}{p-1} \iff t = \frac{nq}{n(p-1) + [p - \alpha(p-1)]q}$$

so that  $\mu \in L^t$  and (67) with the above choice implies  $[M_{p-\alpha(p-1),R}(\mu)]^{1/(p-1)} \in L^q$ . At this point (C1) follows using this fact together with estimate (64). Notice that, recalling that we are assuming  $p \geq 2$ , we have used that

$$t = \frac{nq}{n(p-1) + [p - \alpha(p-1)]q} > 1 \iff q > \frac{n(p-1)}{n - p + \alpha(p-1)} > 1.$$

We have also use the fact with the previous definitions we in fact have that  $t[p - \alpha(p-1)] < n$ .

The assertion in (C2) follows immediately by taking  $\alpha = 0$  in Theorem 11 when  $p = n$ . We indeed recall that by the very definition of BMO spaces we have

$$M_{0,R}^\#(u) \equiv M_R^\#(u) \in L^\infty \iff u \in \text{BMO}. \quad (68)$$

In the case  $2 \leq p < n$ , we recall the following Hölder's type inequality valid in Marcinkiewicz spaces  $\mathcal{M}^t$ ,  $t > 1$ :

$$\|f\|_{L^1(B_\varrho)} \leq \frac{t\varrho^{n(t-1)/t}}{t-1} \|f\|_{\mathcal{M}^t(B_\varrho)}. \quad (69)$$

This implies that in our case  $|\mu|(B_\varrho) \lesssim \varrho^{n-p}$  holds and, by the definition of fractional maximal operator, we conclude that  $M_{p,R}(\mu)$  is locally bounded. At this stage (C2) follows from estimate (44), where we again take  $\alpha = 0$ .

The statement in (C3) is simply an obvious consequence of Theorem 1.

As for (C4), a direct computation using certain characterisation of Lorentz spaces shows that if  $\mu \in L(n/p, 1/(p-1))$  for  $p < n$ , then (21) holds (see for instance [23] for related computations). At this stage (C4) follows by Theorem 1.

In (C5) the first implication follows again by the Hölder type inequality in (69). On the other hand, we observe that the inequality  $|\mu|(B_\varrho) \lesssim \varrho^{n-p+\alpha(p-1)}$  implies that  $M_{p-\alpha(p-1),R}(\mu) \in L^\infty$ , locally, by the very definition of fractional maximal operator in Definition 6, so that (C5) follows by estimate (64).

(C6) is just an obvious consequence of the following imbedding property of Riesz potentials (see [1]):

$$\mathbf{I}_1 : \mathcal{M}_b \rightarrow \mathcal{M}^{n/(n-1)}$$

and of Theorem 3.

Similarly, (C7) follows from Theorem 3 and the classical mapping property

$$\mathbf{I}_1 : L(q, \gamma) \rightarrow L(nq/(n-q), \gamma) \quad \text{for } 1 < q < n \text{ and } \gamma > 0.$$

See also [26, 56].

(C8) is essentially a corollary of Theorem 3 and of the mapping property of Riesz potentials

$$\mathbf{I}_1 : L^{q,\theta} \rightarrow L^{\theta q/(\theta-q),\theta},$$

originally proved by Adams [1] (see also [56] for a localization).

The first implication in (C9) is a consequence of (69) while the second follows applying Theorem 12 with  $\alpha = 0$ . Indeed, assuming that  $|\mu|(B_\varrho) \lesssim \varrho^{n-1}$  implies that  $M_1(\mu)$  is locally bounded, and therefore so is  $M_{0,R}^\#(Du) \equiv M_R^\#(Du)$ . By definition of sharp maximal function this implies that the gradient belongs to BMO, locally; see (68).

(C10) follows from Theorem 8 in an obvious way.

(C11) is again a consequence of (25) and of some basic computation involving the definition of Lorentz norm (see [22, 41]). Indeed, the condition  $\mu \in L(n, 1)$  allows to conclude that (25) holds, again by basic manipulations on Lorentz equivalent norms.

Finally, exactly as for the proof of (C5), (C12) follows again from (69) and estimate (65).  $\square$

Some of the points in the previous corollary are well known results when considering for instance the model equation (15). The theory above allows to extend and embed them in a more general context where results follow in a unified way. Specifically, for (C2) see [55, 64], for (C6) see [7, 8, 20], for (C7) see [19, 31, 56], for (C8) see [56]. We also remark that the previous corollary is just a sample of what is possible to have using the nonlinear potential estimates approach; further spaces, as for instance Lorentz–Morrey or Besov–Morrey spaces are considerable as well (see [55, 56] for relevant definitions). More rearrangement invariant function spaces regularity results via potentials are contained in [11].

## 10 A basic comparison estimate

This section is devoted to the proof of a comparison estimate between a considered solution  $u \in W^{1,p}(\Omega)$  to (17) (notice that we are allowing for measurable coefficients here) and the function  $v \in u + W_0^{1,p}(B_R)$  defined by solving the Dirichlet problem

$$\begin{cases} \operatorname{div} a(x, Dv) = 0 & \text{in } B_R \\ v = u & \text{on } \partial B_R. \end{cases} \quad (70)$$

Here  $B_R \subset \Omega$  denotes a fixed ball with radius  $R$ . The monotonicity properties of the vector field  $a(\cdot)$  described in (18) can be restated and managed more easily via the use of the auxiliary vector field  $V: \mathbb{R}^n \rightarrow \mathbb{R}^n$  defined by

$$V(z) := (|z|^2 + s^2)^{(p-2)/4} z. \quad (71)$$

The number  $s$  is the one defined in (16) and (18). Indeed, we start observing that the following inequality:

$$\frac{|z_1 - z_2|}{c} \leq \frac{|V(z_1) - V(z_2)|}{(|z_1|^2 + |z_2|^2 + s^2)^{(p-2)/4}} \leq c|z_1 - z_2| \quad (72)$$

holds, and is valid for all matrixes  $z_1, z_2 \in \mathbb{R}^n$  that are not simultaneously null (in the case  $s = 0$ ) and for every  $p > 1$ . The constant  $c$  depends only on  $n$  and  $p$ . See [27, 55] for basic properties and for a discussion on the use of the map  $V(\cdot)$  in this context. Since we are assuming  $p \geq 2$ , in particular we have that

$$|z_1 - z_2|^{p/2} \leq c|V(z_1) - V(z_2)|. \quad (73)$$

Combining (18)<sub>2</sub> and (72) yields

$$\frac{1}{c}|V(z_1) - V(z_2)|^2 \leq \langle a(x, z_2) - a(x, z_1), z_2 - z_1 \rangle \quad (74)$$

again for every choice of  $z_1, z_2, x \in \Omega$  and where  $c \equiv c(n, p, v)$ .

The lemmas in this section are already scattered in [21, 41, 55]. The proofs proposed here are anyway different and shorter. We start with a basic, weighted type energy estimate.

**Lemma 1** *Let  $u \in W^{1,p}(\Omega)$  be a solution to (17) under assumptions (18), and let  $v \in u + W_0^{1,p}(B_R)$  be as in (70). Then*

$$\int_{B_R} \frac{|V(Du) - V(Dv)|^2}{(h + |u - v|)^\xi} d\tilde{x} \leq c \frac{h^{1-\xi}}{\xi - 1} |\mu|(B_R) \quad (75)$$

*holds whenever  $h > 0$  and  $\xi > 1$ , where  $c \equiv c(n, p, v) \geq 1$ . The map  $V(\cdot)$  has been defined in (71).*

*Proof* The weak formulation

$$\int_{B_R} \langle a(\tilde{x}, Du) - a(\tilde{x}, Dv), D\varphi \rangle d\tilde{x} = \int_{B_R} \varphi d\mu,$$

is valid whenever  $\varphi \in W_0^{1,p}(B) \cap L^\infty(B)$ , and we use the test functions

$$\eta \equiv \eta_{\pm} := h^{1-\xi} - (h + (u - v)_{\pm})^{1-\xi},$$

which in fact belong to  $W_0^{1,p}(B) \cap L^{\infty}(B)$ . We recall the standard notation

$$(u - v)_+ := \max\{u - v, 0\} \quad \text{and} \quad (u - v)_- := \max\{v - u, 0\}.$$

Then, notice that

$$\begin{aligned} |I_{\pm}| &:= \left| (\xi - 1) \int_{B_R} \frac{\langle a(\tilde{x}, Du) - a(\tilde{x}, Dv), D(u - v)_{\pm} \rangle}{(h + (u - v)_{\pm})^{\xi}} d\tilde{x} \right| \\ &= \left| \int_{B_R} \langle a(\tilde{x}, Du) - a(\tilde{x}, Dv), D\eta_{\pm} \rangle d\tilde{x} \right| \\ &= \left| \int_{B_R} \eta_{\pm} d\mu \right| \leq h^{1-\xi} |\mu|(B_R) \end{aligned}$$

and therefore, using (74), we conclude with

$$(\xi - 1) \int_{B_R} \frac{|V(Du) - V(Dv)|^2}{(h + |u - v|)^{\xi}} d\tilde{x} \leq c(I_+ - I_-) \leq ch^{1-\xi} |\mu|(B_R)$$

so that (75) follows.  $\square$

Lemma 1 and Sobolev embedding theorem in turn imply a first comparison estimate.

**Lemma 2** *Let  $u \in W^{1,p}(\Omega)$  be a solution to (17) under assumptions (18), and let  $v \in u + W_0^{1,p}(B_R)$  be as in (70). Then*

$$\int_{B_R} |Du - Dv|^q d\tilde{x} \leq c_1 \left[ \frac{|\mu|(B_R)}{R^{n-1}} \right]^{q/(p-1)} \quad (76)$$

and therefore

$$\int_{B_R} |u - v|^q d\tilde{x} \leq \tilde{c}_1 \left[ \frac{|\mu|(B_R)}{R^{n-p}} \right]^{q/(p-1)} \quad (77)$$

hold for constants  $c_1, \tilde{c}_1 \equiv c_1, \tilde{c}_1(n, p, v, q)$ , whenever

$$0 < q < \min \left\{ p, \frac{n(p-1)}{n-1} \right\} =: p_m. \quad (78)$$

*Proof* We restrict ourselves to the proof of (76), since this obviously implies (77) via Poincaré's inequality. Moreover, we can restrict ourselves to prove (76) assuming  $q > 1$ , since the result for the remaining values would follow by Hölder's inequality. Now, let us observe that we can always reduce to the situation when there exists  $\xi > 1$  such that

$$\xi q / (p - q) = q^*, \quad (79)$$

where  $q^*$  is the Sobolev conjugate of  $q$ . Indeed, when  $p \leq n$  then  $p_m = n(p - 1)/(n - 1)$  and therefore  $q < p_m$  implies that  $q/(p - q) < q^* = nq/(n - q)$ , so that (79) follows for the corresponding choice of  $\xi$ . When instead  $p > n$  then we first observe that we can prove (76) in the case  $n \leq q < p$ , since (76) for lower values of  $q$  would then follow by Hölder's inequality. Now, since by definition of Sobolev embedding exponent we can take  $q^*$  as large as we please, we again find a number  $\xi > 1$  for which (79) holds. We then plan to apply Lemma 2 with this choice of  $\xi$  and with the following one of  $h$ :

$$h = R \left( \int_{B_R} |Du - Dv|^q d\tilde{x} \right)^{1/q}.$$

Notice we can always assume that  $h > 0$ , otherwise the assertion of Lemma 2 trivializes. By using the definitions in the last two displays, inequalities (73) and (75), and finally Sobolev embedding theorem, we then have

$$\begin{aligned} & \int_{B_R} |Du - Dv|^q d\tilde{x} \\ &= \int_{B_R} \frac{|Du - Dv|^q}{(h + |u - v|)^{\xi q/p}} (h + |u - v|)^{\xi q/p} d\tilde{x} \\ &\leq c \int_{B_R} \frac{|V(Du) - V(Dv)|^{2q/p}}{(h + |u - v|)^{\xi q/p}} (h + |u - v|)^{\xi q/p} d\tilde{x} \\ &\leq c \left( \int_{B_R} \frac{|V(Du) - V(Dv)|^2}{(h + |u - v|)^{\xi}} d\tilde{x} \right)^{q/p} \left( \int_{B_R} (h + |u - v|)^{\xi q/(p-q)} d\tilde{x} \right)^{(p-q)/p} \\ &\leq c \left( \frac{h^{1-\xi}}{\xi - 1} \frac{|\mu|(B_R)}{R^n} \right)^{q/p} \left[ h^{\xi q/p} + \left( \int_{B_R} |u - v|^{\xi q/(p-q)} d\tilde{x} \right)^{(p-q)/p} \right] \\ &\leq c \left( \frac{h^{1-\xi}}{\xi - 1} \frac{|\mu|(B_R)}{R^n} \right)^{q/p} \left[ h^{\xi q/p} + R^{\xi q/p} \left( \int_{B_R} |Du - Dv|^q d\tilde{x} \right)^{\xi/p} \right] \end{aligned}$$

$$\begin{aligned}
&\leq ch^{q/p} \left[ \frac{|\mu|(B_R)}{R^n} \right]^{q/p} \\
&= c \left( \int_{B_R} |Du - Dv|^q d\tilde{x} \right)^{1/p} \left[ \frac{|\mu|(B_R)}{R^{n-1}} \right]^{q/p} \\
&\leq \frac{1}{2} \int_{B_R} |Du - Dv|^q d\tilde{x} + c \left[ \frac{|\mu|(B_R)}{R^{n-1}} \right]^{q/(p-1)}
\end{aligned}$$

so that (76) follows.  $\square$

**Remark 4** By examining the arguments of the previous lemma it is possible to see that the constant  $c_1$  appearing in (76) shows the following natural asymptotic for  $q \rightarrow p_m$ :

$$c_1 = c(n, p, v) \left( \frac{1}{p_m - q} \right)^{q/(p-1)}, \quad (80)$$

where  $c$  is a constant that remains bounded whenever  $p$  varies in a compact subset of  $(1, \infty)$ ; see (79) and the constant appearing in Sobolev–Morrey embedding theorem. The asymptotic in (80) is typical in situations when a borderline estimate in Lebesgue spaces fails, being replaced by an estimate in Marcinkiewicz spaces; recall (58).

## 11 A sequence of comparison estimates

Given a number  $\delta_1 \in (0, 1/4)$  and a ball  $B(x, r) \subset \Omega$ , we define the sequence of shrinking balls

$$B_j := B(x, r_j), \quad r_j = \delta_1^j r, \quad (81)$$

whenever  $j \geq 0$  is an integer. The related comparison solutions  $v_j \in u + W_0^{1,p}(B_j)$  are as in (70), that is

$$\begin{cases} \operatorname{div} a(x, Dv_j) = 0 & \text{in } B_j \\ v_j = u & \text{on } \partial B_j. \end{cases} \quad (82)$$

Then we have

**Lemma 3** *Let  $u \in W^{1,p}(\Omega)$  be a solution to (17) under assumptions (18), and let  $v_{j-1}, v_j$  be as in (82) with  $j \geq 1$ . Suppose further that*

$$\frac{|\mu|(B_{j-1})}{\lambda^{p-1} r_{j-1}^{n-1}} \leq 1 \quad (83)$$

and that the bounds

$$\frac{\lambda}{A} \leq |Dv_{j-1}| \leq A\lambda \quad \text{in } B_j \quad (84)$$



hold for constants  $A \geq 1$  and  $\lambda > 0$ . Then there exists a constant  $c_2$  depending only on  $n, p, v, \delta_1, A$  such that

$$\int_{B_j} |Du - Dv_j| d\tilde{x} \leq c_2 \lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]. \quad (85)$$

*Proof* We start fixing the following quantities:

$$\gamma := \frac{1}{4(p-1)(n+1)} \quad \text{and} \quad \xi := 1 + 2\gamma, \quad (86)$$

in view of the application of Lemmas 1 and 2, that we shall use with exponents  $q$  such that

$$0 < q \leq \xi(p-1) = p-1 + \frac{1}{2(n+1)}.$$

This in particular fixes the constant  $c_1$  from Lemma 2. We also set

$$\bar{v}_{j-1} := \frac{v_{j-1}}{\lambda}, \quad \bar{v}_j := \frac{v_j}{\lambda}. \quad (87)$$

In the rest of the proof constants denoted by  $c$  will only depend on  $n, p, v, \delta_1, A$  and will in general vary from line to line, as usual. We start estimating the term on the left in (85) with the aid of (84) as follows:

$$\begin{aligned} \int_{B_j} |Du - Dv_j| d\tilde{x} &\leq A^{(p-2)(1+\gamma)} \int_{B_j} |D\bar{v}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \\ &\leq c \int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \\ &\quad + c \int_{B_j} |D\bar{v}_j|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x}. \end{aligned} \quad (88)$$

To continue, we estimate the second-last integral appearing in the previous display. For this let us preliminary note that, by using (76) and recalling that  $r_j = \delta_1 r_{j-1}$ , the following estimate holds indeed for the same range of exponents  $q$  working in (76):

$$\begin{aligned} &\int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^q d\tilde{x} \\ &\leq c \lambda^{-q} \int_{B_j} |Dv_j - Du|^q d\tilde{x} + c \lambda^{-q} \frac{|B_{j-1}|}{|B_j|} \int_{B_{j-1}} |Dv_{j-1} - Du|^q d\tilde{x} \end{aligned}$$

$$\begin{aligned}
&\leq c\lambda^{-q} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{q/(p-1)} + c\lambda^{-q} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]^{q/(p-1)} \\
&\leq c \left[ \frac{|\mu|(B_{j-1})}{\lambda^{p-1} r_{j-1}^{n-1}} \right]^{q/(p-1)}. \tag{89}
\end{aligned}$$

Then, appealing to Hölder's inequality together with (76), gives us

$$\begin{aligned}
&\int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \\
&\leq \left( \int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^{(p-1)(1+\gamma)} d\tilde{x} \right)^{(p-2)/(p-1)} \\
&\quad \cdot \left( \int_{B_j} |Du - Dv_j|^{p-1} d\tilde{x} \right)^{1/(p-1)} \\
&\leq c\lambda \left[ \frac{|\mu|(B_{j-1})}{\lambda^{p-1} r_{j-1}^{n-1}} \right]^{[(1+\gamma)(p-2)+1]/(p-1)}. \tag{90}
\end{aligned}$$

But now, as

$$\frac{(1+\gamma)(p-2)+1}{p-1} > 1 \tag{91}$$

then (83) implies

$$\int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \leq c\lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]$$

with  $c \equiv c(n, p, \nu, \delta_1)$ . Combining this last estimate with (88) gives us

$$\begin{aligned}
&\int_{B_j} |Du - Dv_j| d\tilde{x} \\
&\leq c \int_{B_j} |D\bar{v}_j|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} + c\lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]. \tag{92}
\end{aligned}$$

It remains to estimate the first term on the right hand side in the above display. Applying Hölder's inequality, together with (72) and (75), and recalling that  $\xi = 1 + 2\gamma$ , we obtain for any  $h > 0$  that

$$\begin{aligned}
 & \int_{B_j} |D\bar{v}_j|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \\
 & \leq c \int_{B_j} \left[ \lambda^{2-p} \frac{(|Dv_j| + |Du|)^{p-2} |Du - Dv_j|}{(h + |u - v_j|)^\xi} \right]^{1/2} \\
 & \quad \cdot \left[ |D\bar{v}_j|^{(p-2)(1+2\gamma)} (h + |u - v_j|)^\xi \right]^{1/2} d\tilde{x} \\
 & \leq c \int_{B_j} \left[ \lambda^{2-p} \frac{|V(Du) - V(Dv_j)|^2}{(h + |u - v_j|)^\xi} \right]^{1/2} \left[ |D\bar{v}_j|^{(p-2)\xi} (h + |u - v_j|)^\xi \right]^{1/2} d\tilde{x} \\
 & \leq c \left( \lambda^{2-p} \int_{B_j} \frac{|V(Du) - V(Dv_j)|^2}{(h + |u - v_j|)^\xi} d\tilde{x} \right)^{1/2} \\
 & \quad \cdot \left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} (h + |u - v_j|)^\xi d\tilde{x} \right)^{1/2} \\
 & \leq c \sqrt{\lambda^{2-p}} \left[ h^{1-\xi} \frac{|\mu|(B_j)}{r_j^n} \right]^{1/2} \left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} (h + |u - v_j|)^\xi d\tilde{x} \right)^{1/2}.
 \end{aligned} \tag{93}$$

Now we choose

$$h := \left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} |u - v_j|^\xi d\tilde{x} \right)^{1/\xi} + \delta \tag{94}$$

for some small positive  $\delta \in (0, 1)$  to get

$$\begin{aligned}
 & \left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} (h + |u - v_j|)^\xi d\tilde{x} \right)^{1/2} \\
 & \leq 2h^{\xi/2} \left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} d\tilde{x} \right)^{1/2} + 2h^{\xi/2}.
 \end{aligned}$$

Notice that the presence in (94) of the parameter  $\delta$  is aimed at guaranteeing that  $h$  is positive; we shall eventually let  $\delta \rightarrow 0$  at the end of the proof. In the above display, we use (89) and (84) to gain

$$\begin{aligned} \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} d\tilde{x} &\leq c \int_{B_j} |D\bar{v}_{j-1} - D\bar{v}_j|^{(p-2)\xi} d\tilde{x} + c \int_{B_j} |D\bar{v}_{j-1}|^{(p-2)\xi} d\tilde{x} \\ &\leq c \left[ \frac{|\mu|(B_{j-1})}{\lambda^{p-1} r_{j-1}^{n-1}} \right]^{(p-2)\xi/(p-1)} + c A^{(p-2)\xi} \leq c, \end{aligned}$$

so that we conclude with

$$\left( \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} (h + |u - v_j|)^\xi d\tilde{x} \right)^{1/2} \leq ch^{\xi/2},$$

where the constant  $c$  ultimately depends only on  $n, p, v, \delta_1, A$ . Plugging the last estimate in (93) now yields

$$\int_{B_j} |D\bar{v}_j|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \leq c \sqrt{\frac{h}{r_j}} \left[ \frac{|\mu|(B_j)}{\lambda^{p-2} r_j^{n-1}} \right]^{1/2}$$

while applying Young's inequality and recalling again that  $r_j = \delta_1 r_{j-1}$  gives

$$\int_{B_j} |D\bar{v}_j|^{(p-2)(1+\gamma)} |Du - Dv_j| d\tilde{x} \leq \frac{c\lambda^{2-p}}{\varepsilon} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right] + \frac{\varepsilon h}{r_{j-1}} \quad (95)$$

whenever  $\varepsilon \in (0, 1)$ , where  $c \equiv c(n, p, v, \delta_1, A)$  is in particular independent of  $\varepsilon$ . It finally remains to estimate  $h$ , which has been defined in (94). We have

$$\begin{aligned} h &\leq c \left( \int_{B_j} |D\bar{v}_{j-1} - D\bar{v}_j|^{(p-2)\xi} |u - v_j|^\xi d\tilde{x} \right)^{1/\xi} \\ &\quad + c \left( \int_{B_j} |D\bar{v}_{j-1}|^{(p-2)\xi} |u - v_j|^\xi d\tilde{x} \right)^{1/\xi} + \delta =: I_1 + I_2 + \delta. \quad (96) \end{aligned}$$

Using (76) and (89) to estimate  $I_1$  leads to

$$\begin{aligned}
 I_1 &\leq c \left( \int_{B_j} |D\bar{v}_{j-1} - D\bar{v}_j|^{\xi(p-1)} d\tilde{x} \right)^{(p-2)/[\xi(p-1)]} \\
 &\quad \cdot \left( \int_{B_j} |u - v_j|^{\xi(p-1)} d\tilde{x} \right)^{1/[\xi(p-1)]} \\
 &\leq cr_j \left[ \frac{|\mu|(B_{j-1})}{\lambda^{p-1} r_{j-1}^{n-1}} \right]^{(p-2)/(p-1)} \left( \int_{B_j} |Du - Dv_j|^{\xi(p-1)} d\tilde{x} \right)^{1/[\xi(p-1)]} \\
 &\leq cr_{j-1} \lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]. \tag{97}
 \end{aligned}$$

In turn we bound  $I_2$  appealing to (76) and using (84) repeatedly as follows:

$$\begin{aligned}
 I_2 &\leq cA^{p-2} \left( \int_{B_j} |u - v_j|^{\xi} d\tilde{x} \right)^{1/\xi} \\
 &\leq c \left( \int_{B_j} |u - v_j|^{n/(n-1)} d\tilde{x} \right)^{(n-1)/n} \\
 &\leq cr_j \int_{B_j} |Du - Dv_j| d\tilde{x} \\
 &\leq cA^{(p-2)\xi} r_{j-1} \int_{B_j} |D\bar{v}_{j-1}|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} \\
 &\leq cr_{j-1} \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} \\
 &\quad + cr_{j-1} \int_{B_j} |D\bar{v}_{j-1} - D\bar{v}_j|^{(p-2)\xi} |Du - Dv_j| d\tilde{x}.
 \end{aligned}$$

Similarly to (90) we find

$$\int_{B_j} |D\bar{v}_j - D\bar{v}_{j-1}|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} \leq c\lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]$$

so that

$$I_2 \leq cr_{j-1} \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} + cr_{j-1} \lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right].$$

Combining this last estimate with (97) and eventually with (96) yields

$$\frac{h}{r_{j-1}} \leq c_* \int_{B_j} |D\bar{v}_j|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} + c_* \lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right] + \frac{\delta}{r_{j-1}},$$

where  $c_*$  depends only on  $n, p, v, L, A, \delta_1$ . Plugging the inequality in the above display in (95), choosing  $\varepsilon = 1/(2c_*)$  and reabsorbing terms leads to

$$\int_{B_j} |D\bar{v}_j|^{(p-2)\xi} |Du - Dv_j| d\tilde{x} \leq c \lambda^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right] + \frac{c\delta}{r_{j-1}}.$$

Now (85) follows using this last inequality in combination with the one in display (92), and finally letting  $\delta \rightarrow 0$ .  $\square$

## 12 Proof of Theorems 8 and 10

The proof of Theorem 10 falls in six steps, going through Sects. 12.1–12.6 respectively. It requires the use of several different tools, and ultimately relies on a delicate iteration technique. Finally, in Sect. 12.7, we shall briefly show how to get Theorem 8 from Theorem 10.

In the following, given a ball  $B \subset \Omega$ , we define the excess of a vector field  $f \in L^1(B, \mathbb{R}^k)$  over  $B$  as

$$E(f, B) := \int_B |f - (f)_B| d\tilde{x}.$$

This functional, roughly speaking, provides an integral measure of the oscillations of the vector field  $f$  in the ball  $B$ . An elementary property of the excess is given by the following inequality:

$$\int_B |f - (f)_B| d\tilde{x} \leq 2 \int_B |f - \gamma| d\tilde{x} \quad \text{for every } \gamma \in \mathbb{R}^n. \quad (98)$$

We remark that here and in the following we shall use the comparison estimates presented in Sects. 10 and 11 that in fact applies in particular to the case of the solutions considered in Theorem 10.

## 12.1 A density property of $a$ -harmonic functions

In this section we shall consider a solution  $v \in W_{\text{loc}}^{1,p}(B) \cap W^{1,1}(B)$  to the equation

$$\operatorname{div} a(Dv) = 0 \quad \text{in } B, \quad (99)$$

where  $B \subset \mathbb{R}^n$  is a given ball. We here restate in a suitable form the basic regularity properties a priori estimates for solutions to equation in (99); proofs can be for instance found in [18, 21, 51, 55]. The first is the classical gradient  $L^\infty - L^1$  bound

$$\sup_{\gamma B} |Dv| + s \leq \frac{c_l}{(1-\gamma)^n} \int_B (|Dv| + s) d\tilde{x}, \quad (100)$$

which is valid whenever  $\gamma \in (0, 1)$  and for constant  $c_l$  depending only on  $n, p, v, L$ . Next, we restate in suitable way the gradient Hölder continuity estimate included in display (35); more precisely, we select  $\beta \equiv \beta(n, p, v, L) \in (0, \alpha_M)$  such that

$$|Dv(x_1) - Dv(x_2)| \leq c_h \sigma^\beta \int_B (|Dv| + s) d\tilde{x} \quad (101)$$

holds whenever  $\sigma B \subset B$  is a ball concentric to  $B$  with  $\sigma \in (0, 1/2)$ , such that  $x_1, x_2 \in \sigma B$ , where again  $c_h \geq 1$  depends only on  $n, p, v, L$ . Based on the previous regularity estimates, we can prove the following result, that roughly tells that something that happens in average, happens actually everywhere, and in a precisely quantifiable way.

**Proposition 2** (Density improvement) *Let  $v \in W_{\text{loc}}^{1,p}(B) \cap W^{1,1}(B)$  be a solution to (99) such that*

$$\frac{\lambda}{\Gamma} \leq \int_{\sigma^k B} |Dv| d\tilde{x} \quad \text{and} \quad \sup_{B/2} (|Dv| + s) \leq \Gamma \lambda \quad (102)$$

*hold for some integer  $k \geq 1$  and for numbers  $\Gamma \geq 1, \lambda > 0$  and  $\sigma \in (0, 1/4)$  such that*

$$0 < \sigma \leq \left( \frac{1}{4c_h \Gamma^2} \right)^{1/\beta}, \quad (103)$$

*where  $\beta$  and  $c_h$  are the constants appearing in (101). Then*

$$\frac{\lambda}{4\Gamma} \leq |Dv| \quad \text{holds in } \sigma B.$$

*Proof* The first inequality in (102) implies that there exists a point  $x_0 \in \sigma^k B$  such that

$$|Dv(x_0)| > \frac{\lambda}{2\Gamma}.$$

On the other hand, as (101) holds, the second inequality in (102) gives

$$|Dv(x) - Dv(x_0)| \leq c_h \Gamma \lambda \sigma^\beta$$

whenever  $x \in \sigma B$ , the last two inequalities and (103) then give

$$|Dv(x)| \geq |Dv(x_0)| - |Dv(x) - Dv(x_0)| \geq \frac{\lambda}{2\Gamma} - \frac{\lambda}{4\Gamma} = \frac{\lambda}{4\Gamma}$$

for all  $x \in \sigma B$ . □

Finally, we recall two decay estimates that hold for the excess functional of the gradient and of the solutions. They are again related to the Hölder continuity of the gradient of solutions to (99) and of the solutions themselves, in the case we are considering homogeneous equations with measurable coefficients of the type

$$\operatorname{div} a(x, Dv) = 0 \quad \text{in } B. \quad (104)$$

We now connect two basic regularity results for solutions to (99) and (104), respectively. They are bound to provide

**Theorem 17** [19, 21, 47] *Let  $v \in W_{\text{loc}}^{1,p}(B) \cap W^{1,1}(B)$  be a weak solution to (99) under the assumptions (16). There exist constants  $\beta \in (0, 1]$  and  $c_d \geq 1$ , both depending only on  $n, p, \nu, L$ , such that the estimate*

$$\int_{\sigma B} |Dv - (Dv)_{\sigma B}| d\tilde{x} \leq c_d \sigma^\beta \int_B |Dv - (Dv)_B| d\tilde{x} \quad (105)$$

*holds whenever  $\sigma B \subset B$  are concentric balls.*

**Remark 5** The exponent  $\beta$  appearing in (105) can be chosen arbitrarily close to the exponent  $\alpha_M$  appearing in (34). As a matter of fact, estimate (35) is in fact obtained as a corollary of the one in (105) via the standard Campanato's integral characterisation of Hölder continuous functions [10]. The exponent  $\alpha_M$  considered in Theorem 9 is indeed nothing but the sup of the numbers  $\beta$  considerable in Theorem 17. Explicit, though not optimal, estimates on these numbers via the structure parameters  $n, p, \nu, L$ , are retrievable tracking the dependence of the constants in the proofs in [19, 21, 47].

**Theorem 18** *Let  $v \in W_{\text{loc}}^{1,p}(B) \cap W^{1,1}(B)$  be a weak solution to (104) under the assumptions (18), and where the vector field  $a: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  has measurable dependence on  $x$ . There exist constants  $\tilde{\beta} \in (0, 1]$  and  $\tilde{c}_d \geq 1$ , both depending only on  $n, p, \nu, L$ , such that the estimate*



$$\int_{\sigma B} |v - (v)_{\sigma B}| d\tilde{x} \leq \tilde{c}_d \sigma^{\tilde{\beta}} \int_B |v - (v)_B| d\tilde{x} + \tilde{c}_d \sigma^{\tilde{\beta}} \tilde{r} s \quad (106)$$

holds whenever  $\sigma B \subset B$  are concentric balls, where  $\tilde{r}$  is the radius of  $B$ .

*Proof* The proof combines a few well-known facts in the regularity theory of the equation considered. Let us recall the following Caccioppoli type inequality, that states there exists a constant  $c \equiv c(n, p, v, L)$  such that

$$\int_{B/2} (|Dv| + s) d\tilde{x} \leq \frac{c}{|B|^{1/n}} \int_B |v - (v)_B| d\tilde{x} + cs \quad (107)$$

holds; see [40, Proposition 4.1]. On the one hand, the following gradient decay estimate holds whenever  $\sigma \in (0, 1/2]$  (see for instance [55, Lemma 3.3]):

$$\int_{\sigma B} (|Dv| + s) d\tilde{x} \leq c\sigma^{-1+\tilde{\beta}} \int_{B/2} (|Dv| + s) d\tilde{x} \quad (108)$$

where  $\tilde{\beta} < \alpha_m$  and  $c \equiv c(n, p, v, L, \tilde{\beta})$ . Both the inequalities in the last displays are typical in the literature (see for instance [25, Chapter 7]) when using  $L^p$ -norms instead of the  $L^1$ -ones. Combining them then yields

$$\begin{aligned} \int_{\sigma B} |v - (v)_{\sigma B}| d\tilde{x} &\leq c\sigma |B|^{1/n} \int_{\sigma B} |Dv| d\tilde{x} \\ &\leq c\sigma^{\tilde{\beta}} |B|^{1/n} \int_{B/2} (|Dv| + s) d\tilde{x} \\ &\leq c\sigma^{\tilde{\beta}} \int_B |v - (v)_B| d\tilde{x} + c\sigma^{\tilde{\beta}} |B|^{1/n} s \end{aligned}$$

that proves (107) in the case  $\sigma \in (0, 1/2]$ . On the other hand, the remaining case  $\sigma \in (1/2, 1]$  is trivial by using (98).  $\square$

*Remark 6* Exactly as in Remark 5, the exponent  $\tilde{\beta}$  appearing in Theorem 18 can be actually taken as close to  $\alpha_m$ —appearing in (30)–(31)—as we please. As a matter of fact estimates (31) and (106) are equivalent by noticing that if  $v$  solves then so does  $v - k$ , whenever  $k$  is a real number. Again, the exponent  $\alpha_m$  used in Theorem 6 is actually defined as the sup of the numbers  $\beta$  for which (106) works. Estimates for such numbers are available in terms on the structure parameters  $n, p, v, L$ .

## 12.2 Setting of the constants

In the following all the balls considered will be centered at the point  $x \in \Omega$ , and we start form a ball  $B_R \equiv B(x, R) \subset \Omega$  as in the statement of the theorem. We introduce  $\lambda_M$  as

$$\lambda_M := HR^{1-\alpha} \int_{B_R} (|Du| + s) d\tilde{x} + H \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) \right]^{1/(p-1)}, \quad (109)$$

and fix the constant  $H \geq 1$  in a few lines (see (114) below), in a way that makes it depending only on  $n, p, \nu, L$ ; we will then prove that

$$M_{1-\alpha, R}(Du)(x) \leq c\lambda_M \quad (110)$$

holds uniformly in  $\alpha \in [0, 1]$ , for yet another constant  $c \equiv c(n, p, \nu, L)$ . Clearly, we may assume without loss of generality that  $\lambda_M > 0$ , otherwise there is nothing to prove. Let us now fix a few constants that will be relevant in the following; the constants  $c_l, c_h, c_d$  have been defined in (100), (101) and Theorem 17, respectively, while  $c_1$  has been introduced in Lemma 2 (that here will be used with  $q = 1$ ); all these constants depend only on  $n, p, \nu, L$ . Again,  $\beta \equiv \beta(n, p, \nu, L) \in (0, 1)$  is the exponent appearing in (101) and (105). Then we fix

$$\delta_1 := \left( \frac{1}{10^{8n} c_l^2 c_d c_h} \right)^{1/\beta}. \quad (111)$$

It follows that also  $\delta_1$  depends only on  $n, p, \nu, L$ . With such a choice of  $\delta_1$  we consider the balls in (81), that is

$$B_j := B(x, r_j), \quad r_j = \delta_1^j r, \quad r := R/2,$$

and the related comparison solutions  $v_j \in u + W_0^{1,p}(B_j)$  in (82), that is

$$\begin{cases} \operatorname{div} a(Dv_j) = 0 & \text{in } B_j \\ v_j = u & \text{on } \partial B_j. \end{cases} \quad (112)$$

Again, with  $\delta_1$  now being fixed we determine the constant  $c_2$  from Lemma 3 as follows:

$$c_2 := c_2(n, p, \nu, \delta_1, A) \quad \text{with} \quad A := 10^{4n} c_l. \quad (113)$$

The following inclusions hold for every  $j \geq 0$ , and they will be used throughout the proof:  $\delta_1 B_j = B_{j+1} \subset B_j \subset (1/4)B_j \subset B_j$ . We finally set

$$H := 10^8 c_1 c_2 c_d \delta_1^{-5n} \quad (114)$$

and again notice that  $H$  only depends on  $n, p, \nu, L$ . Now, let us set

$$\kappa := \min \left\{ 1, \delta_1^{n-p+\alpha(p-1)} \right\} \quad \text{and} \quad \tilde{r}_j := \frac{r_j + r_{j+1}}{2}$$

for  $j \geq 0$ , and observe that

$$\begin{aligned}
 \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) &= \int_0^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} \\
 &= \sum_{j=0}^{\infty} \int_{\tilde{r}_j}^{r_j} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} + \int_{8r/7}^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} \\
 &\quad + \sum_{j=0}^{\infty} \int_{r_{j+1}}^{\tilde{r}_j} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} + \int_r^{8r/7} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} \\
 &\geq \sum_{j=0}^{\infty} \int_{\tilde{r}_j}^{r_j} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} + \int_{8r/7}^{2r} \frac{|\mu|(B_\varrho)}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} \\
 &\geq \kappa \sum_{j=0}^{\infty} \frac{|\mu|(\tilde{B}_{j+1})}{r_{j+1}^{n-p+\alpha(p-1)}} \int_{\tilde{r}_j}^{r_j} \frac{d\varrho}{\varrho} + \frac{|\mu|(\tilde{B}_0)}{2^{n-1} r^{n-p+\alpha(p-1)}} \int_{8r/7}^{2r} \frac{d\varrho}{\varrho} \\
 &\geq \delta_1^{n-1} \log\left(\frac{2}{1+\delta_1}\right) \sum_{j=0}^{\infty} \frac{|\mu|(\tilde{B}_{j+1})}{r_{j+1}^{n-p+\alpha(p-1)}} + \frac{\log(7/4)}{2^{n-1}} \left[ \frac{|\mu|(\tilde{B}_0)}{r_0^{n-p+\alpha(p-1)}} \right] \\
 &\geq \delta_1^n \sum_{j=0}^{\infty} \frac{|\mu|(\tilde{B}_j)}{r_j^{n-p+\alpha(p-1)}}. \tag{115}
 \end{aligned}$$

Therefore, by (109) and the choice in (114) it follows that

$$10^8 c_1 c_2 c_d \delta_1^{-4n} \left[ \sum_{j=0}^{\infty} \frac{|\mu|(\tilde{B}_j)}{r_j^{n-p+\alpha(p-1)}} \right]^{1/(p-1)} \leq \lambda_M. \tag{116}$$

Of course we are using constants like  $10^8$  to emphasize the fact that in certain places of the proof what it matters is to take large/small quantities. In the rest of the proof we shall denote, for every  $j \geq 0$ ,

$$E_j := E(Du, B_j) \quad \text{and} \quad a_j := |(Du)_{B_j}| \tag{117}$$

and

$$\lambda_j := r_j^{\alpha-1} \lambda_M. \tag{118}$$

### 12.3 Iterating quantities

The proof is based on the fact that certain quantities, related to the fractional maximal operator we are considering, stay bounded. Specifically, we shall prove that the following inequality holds:

$$\sup_{j \geq 1} r_j^{1-\alpha} \int_{B_j} (|Du| + s) d\tilde{x} \leq \lambda_M. \quad (119)$$

Let us briefly recall that the last inequality implies the one in display (110) with  $c = (2/\delta_1)^n$ . Indeed, for  $0 < \varrho \leq r_1$  we find  $i \geq 1$  such that  $r_{i+1} = \delta_1 r_i < \varrho \leq r_i$ , therefore

$$\varrho^{1-\alpha} \int_{B_\varrho} (|Du| + s) d\tilde{x} \leq r_i^{1-\alpha} \left(\frac{r_i}{\varrho}\right)^n \int_{B_i} (|Du| + s) d\tilde{x} \leq \frac{\lambda_M}{\delta_1^n}.$$

On the other hand, if  $r_1 < \varrho \leq R$ , recalling that  $r = R/2$ , then we similarly have

$$\varrho^{1-\alpha} \int_{B_\varrho} (|Du| + s) d\tilde{x} \leq R^{1-\alpha} \left(\frac{R}{\varrho}\right)^n \int_{B_R} (|Du| + s) d\tilde{x} \leq \frac{2^n \lambda_M}{\delta_1^n}$$

where this time we have used directly (109). All in all (110) follows with  $c = (2/\delta_1)^n$ . Let us now define

$$C_j := r_j^{1-\alpha} \int_{B_j} |Du| d\tilde{x} + r_{j-1}^{1-\alpha} \int_{B_{j-1}} |Du| d\tilde{x} + \delta_1^{-n} r_j^{1-\alpha} \int_{B_j} |Du - (Du)_{B_j}| d\tilde{x}$$

whenever  $j \geq 1$ . Notice that, since  $r_{j-1}^{1-\alpha} \leq R^{1-\alpha}$  for every  $j \geq 1$ , then the definitions in (109) and (114) imply that

$$C_1 \leq 4\delta_1^{-3n} R^{1-\alpha} \int_{B_R} (|Du| + s) d\tilde{x} \leq \frac{\lambda_M}{100}. \quad (120)$$

Notice that the last inequality implies, in particular, that

$$R^{1-\alpha} s \leq \frac{\lambda_M}{100}. \quad (121)$$

The rest of the proof is now devoted to establish (119). This will be done using induction on certain *interaction chains* defined in the next step.

## 12.4 Iteration chains

We consider the set  $\mathcal{L}$  defined by

$$\mathcal{L} := \left\{ j \in \mathbb{N} : C_j \leq \frac{\lambda_M}{100} \right\}, \quad (122)$$

and, accordingly, for  $k \geq 1$ , we then define the set

$$\mathcal{C}_i^k = \{j \in \mathbb{N} : i \leq j \leq i+k, i \in \mathcal{L}, i+k+1 \in \mathcal{L}, j \notin \mathcal{L} \text{ if } j > i\} \quad (123)$$

and call it *maximal iteration chain* of length  $k$ , starting at  $i$ . In other words, we have  $\mathcal{C}_i^k = \{i, \dots, i+k\}$  and each element of  $\mathcal{C}_i^k$  but  $i$  lies outside of  $\mathcal{L}$ ;  $\mathcal{C}_i^k$  is maximal in the sense that there cannot be another set of the same type properly containing it. Obviously, such sets do not exist when  $\mathcal{L} = \mathbb{N}$ . In the same way we define

$$\mathcal{C}_i^\infty = \{j \in \mathbb{N} : i \leq j < \infty, i \in \mathcal{L}, j \notin \mathcal{L} \text{ if } j > i\} \quad (124)$$

as *the infinite maximal chain* starting at  $i$ . Notice that, in every case, the smallest element of such a chain always belongs to  $\mathcal{L}$ , being then the only one of the chain to have such a property. Now, observe that if  $\mathcal{L} = \mathbb{N} \setminus \{0\}$ , then we are finished as in this case we have  $C_j \leq \lambda_M/100$  for every  $j \geq 1$  and therefore

$$r_j^{1-\alpha} \int_{B_j} (|Du| + s) d\tilde{x} \leq C_j + R^{1-\alpha}s \leq \frac{\lambda_M}{100} + \frac{\lambda_M}{100} \leq \lambda_M \quad (125)$$

where we also used (121). Let us then assume that  $\mathcal{L} \neq \mathbb{N} \setminus \{0\}$ . Since anyway  $1 \in \mathcal{L}$  by (120), then there must be at least one iteration chain. Let us therefore consider one of these,  $\mathcal{C}_i^k$ ,  $k \leq \infty$ . We shall use finite induction to show that

$$\sup_{i \leq j \leq i+k} r_j^{1-\alpha} \int_{B_j} (|Du| + s) d\tilde{x} \leq \lambda_M. \quad (126)$$

This, due to the fact that  $\mathcal{C}_i^k$  is arbitrary, will finally prove (119) since if a certain index  $j$  does not belong to  $\mathcal{L}$ , then it must belong to some chain. On the other hand, if  $j$  does not belong to any chain then it belongs to  $\mathcal{L}$  and the corresponding inequality in (119) is automatically verified by the definition of  $C_j$  and (125). We are now reduced to show the validity of (126) for any iteration chain  $\mathcal{C}_i^k$ .

## 12.5 Density and decay properties along iteration chains

The idea is now that along iteration chains, a condition of the type  $C_j \geq \lambda_M/100$  guarantees that the equation becomes nondegenerate and can be in a sense linearized. The consequence is that the excess functional of the gradient decays in a way that resembles the one of solutions to the Poisson equation. To implement this we need

to get a few density estimates. Let us preliminary prove a *cheap decay estimate*, that actually holds also outside the context of the present proof, but whose assumptions are obviously satisfied here by (111).

**Lemma 4** (Cheap decay estimate) *Let the maps  $v_j \in u + W_0^{1,p}(B_j)$  be defined in (112) and the balls  $\{B_j\}$  be defined in (81). Assume that the number  $\delta_1$  satisfies*

$$\delta_1 \leq \left( \frac{1}{10^7 c_d} \right)^{1/\beta}. \quad (127)$$

*Then the following estimate holds for every  $j \geq 0$  and for a constant depending only on  $n, p, v, L$  but not on  $\alpha \in [0, 1]$ :*

$$r_{j+1}^{1-\alpha} E_{j+1} \leq \frac{1}{10^6} r_j^{1-\alpha} E_j + 8c_1 c_d \delta_1^{-n} \left[ \frac{|\mu|(B_j)}{r_j^{n-p+\alpha(p-1)}} \right]^{1/(p-1)}. \quad (128)$$

*Here the numbers  $\{E_j\}$  have been defined in (117), and the constants  $c_1, c_d$  have been introduced in (76) and (105), respectively.*

*Proof* By using, in order, triangle inequality, Lemma 2 and Theorem 17 with  $v \equiv v_j$ , and also using (98) repeatedly, we have

$$\begin{aligned} r_{j+1}^{1-\alpha} E_{j+1} &\leq 2r_{j+1}^{1-\alpha} \int_{B_{j+1}} |Du - (Dv_j)_{B_{j+1}}| d\tilde{x} \\ &\leq 2r_{j+1}^{1-\alpha} \int_{B_{j+1}} |Dv_j - (Dv_j)_{B_{j+1}}| d\tilde{x} \\ &\quad + 2\delta_1^{-n} r_{j+1}^{1-\alpha} \int_{B_j} |Dv_j - Du| d\tilde{x} \\ &\leq 2c_d \delta_1^\beta r_j^{1-\alpha} \int_{B_j} |Dv_j - (Dv_j)_{B_j}| d\tilde{x} \\ &\quad + 2c_1 \delta_1^{-n} r_{j+1}^{1-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \\ &\leq 4c_d \delta_1^\beta r_j^{1-\alpha} \int_{B_j} |Du - (Du)_{B_j}| d\tilde{x} \\ &\quad + 8c_1 c_d \delta_1^{-n} r_j^{1-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \\ &= 4c_d \delta_1^\beta r_j^{1-\alpha} E_j + 8c_1 c_d \delta_1^{-n} r_j^{1-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \end{aligned}$$

so that (4) follows by (127) and noticing that

$$r_j^{1-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} = \left[ \frac{|\mu|(B_j)}{r_j^{n-p+\alpha(p-1)}} \right]^{1/(p-1)}. \quad (129)$$

□

The following lemma is an easy consequence of the previous one:

**Lemma 5** *Let  $j \geq 0$  be an integer; if*

$$\delta_1^{-n} r_j^{1-\alpha} E_j \leq \frac{\lambda_M}{100}$$

where  $\lambda_M > 0$  is defined in (109), then it also holds that

$$\delta_1^{-n} r_{j+1}^{1-\alpha} E_{j+1} \leq \frac{\lambda_M}{10^5}. \quad (130)$$

*Proof* Using (128) we have

$$\delta_1^{-n} r_{j+1}^{1-\alpha} E_{j+1} \leq \frac{\delta_1^{-n}}{10^6} r_j^{1-\alpha} E_j + 8c_1 c_d \delta_1^{-2n} \left[ \frac{|\mu|(B_j)}{r_j^{n-p+\alpha(p-1)}} \right]^{1/(p-1)} \leq \frac{\lambda_M}{10^5}$$

where in the last estimate we have employed also (116). □

The cheap decay estimate for the excess functional of Lemma 4 is not sufficient to get gradient estimates via linear Riesz potentials. Indeed, it only implies Wolff potential estimates as in (23), as shown in [21]. Anyway, by using it together with the density properties of Proposition 2 and Lemma 3, we deduce another, better decay estimate, which is this time implying linear potentials estimates.

**Lemma 6** (Linearized decay estimate) *Let  $j \geq 1$  be an integer; if*

$$r_{j-1}^{1-\alpha} \int_{B_{j-1}} (|Du| + s) d\tilde{x} \leq \lambda_M \quad (131)$$

holds and if it happens that

$$\delta_1^{-n} r_j^{1-\alpha} E_j \leq \frac{\lambda_M}{100} \quad \text{and} \quad C_{j+1} \geq \frac{\lambda_M}{100} \quad (132)$$

then we have

$$r_{j+1}^{1-\alpha} E_{j+1} \leq \frac{1}{4} r_j^{1-\alpha} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right], \quad (133)$$

where  $c_2$  depends only on  $n, p, v, L$  and has been defined in (113).

*Proof* Using triangle inequality, the second inequality in (132) and Lemma 2 we estimate

$$\begin{aligned}
 & \delta_1^{-n} r_{j+1}^{1-\alpha} E_{j+1} + \sum_{k=j,j+1} r_k^{1-\alpha} \int_{B_k} |Dv_{j-1}| d\tilde{x} \\
 & \geq C_{j+1} - \sum_{k=j,j+1} r_k^{1-\alpha} \int_{B_k} |Du - Dv_{j-1}| d\tilde{x} \\
 & \geq \frac{\lambda_M}{100} - \sum_{k=j,j+1} r_k^{1-\alpha} \int_{B_k} |Du - Dv_{j-1}| d\tilde{x} \\
 & \geq \frac{\lambda_M}{100} - 2\delta_1^{-2n} r_{j-1}^{1-\alpha} \int_{B_{j-1}} |Du - Dv_{j-1}| d\tilde{x} \\
 & \geq \frac{\lambda_M}{100} - 2\delta_1^{-2n} c_1 r_{j-1}^{1-\alpha} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]^{1/(p-1)} \\
 & \geq \frac{\lambda_M}{100} - 2\delta_1^{-2n} c_1 \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right]^{1/(p-1)} \\
 & \geq \frac{\lambda_M}{200}
 \end{aligned}$$

where we have used (129) and (116) to perform the last estimations; we have also used that  $r_k^{1-\alpha} \leq r_{j-1}^{1-\alpha}$ . The previous inequality and (130) then give

$$\frac{\lambda_M}{200} \leq \sum_{k=j,j+1} r_k^{1-\alpha} \int_{B_k} |Dv_{j-1}| d\tilde{x} + \frac{\lambda_M}{10^5}$$

so that we conclude with

$$\max_{k=j,j+1} r_k^{1-\alpha} \int_{B_k} |Dv_{j-1}| d\tilde{x} \geq \frac{\lambda_M}{10^3}$$

and in any case, since  $r_k^{1-\alpha} \leq r_{j-1}^{1-\alpha}$ , we have that

$$\max_{k=j,j+1} \int_{B_k} |Dv_{j-1}| d\tilde{x} \geq \frac{r_{j-1}^{\alpha-1} \lambda_M}{10^3} = \frac{\lambda_{j-1}}{10^3} \geq \frac{\lambda_{j-1}}{10^3 c_l} \quad (134)$$

holds, where the numbers  $\lambda_j$  have been defined in (118). On the other hand, observe that using Lemma 2 and (131) we have



$$\begin{aligned}
r_{j-1}^{1-\alpha} \int_{B_{j-1}} (|Dv_{j-1}| + s) d\tilde{x} &\leq r_{j-1}^{1-\alpha} \int_{B_{j-1}} (|Du| + s) d\tilde{x} \\
&\quad + c_l r_{j-1}^{1-\alpha} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]^{1/(p-1)} \\
&\leq \lambda_M + c_l \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right]^{1/(p-1)}
\end{aligned}$$

and taking again (116) into account we have

$$\int_{B_{j-1}} (|Dv_{j-1}| + s) d\tilde{x} \leq 2r_{j-1}^{\alpha-1} \lambda_M.$$

Applying estimate (100) then yields

$$(|Dv_{j-1}| + s) \leq 2^{n+1} c_l r_{j-1}^{\alpha-1} \lambda_M = 2^{n+1} c_l \lambda_{j-1} \quad \text{in } (1/2)B_{j-1}. \quad (135)$$

We are now in position to apply the density improvement results from Proposition 2; indeed conditions in (102) are satisfied with the choices  $v \equiv v_{j-1}$ ,  $B \equiv B_{j-1}$ ,  $\sigma = \delta_1$ ,  $k = 1$  or  $k = 2$  (by (134)),  $\lambda \equiv \lambda_{j-1}$  and  $\Gamma = 10^{3n} c_l$ ; observe that this is possible by the choice of  $\delta_1$  made in (111). Notice that (103) is satisfied due to the choice of  $\delta_1$  in (111). Proposition 2 now yields

$$\frac{\lambda_{j-1}}{10^{4n} c_l} \leq |Dv_{j-1}| \quad \text{in } B_j = \delta_1 B_{j-1}.$$

The previous inequality and (135) summarise in the following:

$$\frac{\lambda_{j-1}}{A} \leq |Dv_{j-1}| \leq A \lambda_{j-1} \quad \text{in } B_j, \quad A = 10^{4n} c_l.$$

On the other hand, notice that again by (116) we have

$$\frac{|\mu|(B_{j-1})}{\lambda_{j-1}^{p-1} r_{j-1}^{n-1}} = \frac{|\mu|(B_{j-1})}{\lambda_M^{p-1} r_{j-1}^{n-p+\alpha(p-1)}} \leq 1.$$

We are thus in position to apply Lemma 3 thereby obtaining

$$\begin{aligned}
r_j^{1-\alpha} \int_{B_j} |Du - Dv_j| d\tilde{x} &\leq c_2 r_{j-1}^{1-\alpha} \lambda_{j-1}^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right] \\
&= c_2 \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right],
\end{aligned}$$

where the constant  $c_2$  has been fixed in (113). With the last comparison estimate we can conclude as in Lemma 4, that is using Theorem 17 and the inequality in the last display as follows:

$$\begin{aligned}
 r_{j+1}^{1-\alpha} E_{j+1} &\leq 2r_{j+1}^{1-\alpha} \int_{B_{j+1}} |Du - (Dv_j)_{B_{j+1}}| d\tilde{x} \\
 &\leq 2r_{j+1}^{1-\alpha} \int_{B_{j+1}} |Dv_j - (Dv_j)_{B_{j+1}}| d\tilde{x} + 2\delta_1^{-n} r_j^{1-\alpha} \int_{B_j} |Du - Dv_j| d\tilde{x} \\
 &\leq 2c_d \delta_1^\beta r_j^{1-\alpha} \int_{B_j} |Dv_j - (Dv_j)_{B_j}| d\tilde{x} + 2c_2 \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right] \\
 &\leq 4c_d \delta_1^\beta r_j^{1-\alpha} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right] \quad (136)
 \end{aligned}$$

and the proof of (133) is complete recalling that  $4c_d \delta_1^\beta \leq 1/4$  by (111).  $\square$

## 12.6 Iteration and conclusion

Here we finally prove that (126) holds for any iteration chain  $\mathcal{C}_i^k$ , thereby concluding the proof; we remark that by construction any chain  $\mathcal{C}_i^k$  is such that  $i \geq 1$  since the numbers  $C_j$  are defined only for  $j \geq 1$ . We shall first prove by induction that

$$\begin{cases} r_j^{1-\alpha} \int_{B_j} (|Du| + s) d\tilde{x} \leq \lambda_M \\ \delta_1^{-n} r_j^{1-\alpha} E_j \leq \frac{\lambda_M}{100} \\ r_{j+1}^{1-\alpha} E_{j+1} \leq \frac{1}{4} r_j^{1-\alpha} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-p+\alpha(p-1)}} \right] \end{cases} \quad (137)$$

holds for every  $j \in \{i, \dots, i+k-1\}$ . Notice that by definition of iteration chains we always have  $k \geq 1$ . Moreover, notice also that we are going to use finite induction when  $k < \infty$ , that is when the length of the iteration chain is finite; when  $k = \infty$  we are simply going to prove (137) for every  $j \geq i$ . We start by considering the case  $j = i$ , which is the induction basis. This follows by the very definitions of  $\mathcal{C}_i^k$  and  $C_i$  as far as (137)<sub>2</sub> is concerned. As for (137)<sub>1</sub>, notice that using (121) we have

$$r_i^{1-\alpha} \int_{B_i} (|Du| + s) d\tilde{x} \leq C_i + R^{1-\alpha} s \leq \lambda_M.$$

Exactly in the same way we also have

$$r_{i-1}^{1-\alpha} \int_{B_{i-1}} (|Du| + s) d\tilde{x} \leq C_i + R^{1-\alpha} s \leq \lambda_M.$$

This last inequality allows to get (137)<sub>3</sub>. Indeed, we need Lemma 6 with  $j = i$ ; this applies by the inequality in the last display and since by definition of iteration chain  $C_i^k$  we have  $C_{i+1} \geq \lambda_M/100$ ; note that here we are also applying (137)<sub>2</sub>. Next, we proceed verifying the induction step; notice that this case occurs only when  $k \geq 2$  while when  $k = 1$  the first step already concludes the proof of (137) since  $\{i\} = \{i, \dots, i+k-1\}$ . We assume that (137) holds for all indexes  $j \in \{i, \dots, h\}$  with  $h < i+k-1$ , and then we prove it for the index  $h+1$ . By Lemma 5 and (137)<sub>2</sub> (for  $j = h$ ), we immediately have

$$\delta_1^{-n} r_{h+1}^{1-\alpha} E_{h+1} \leq \frac{\lambda_M}{100}. \quad (138)$$

Next, we are going to prove that

$$r_{h+1}^{1-\alpha} \int_{B_{h+1}} (|Du| + s) d\tilde{x} \leq \lambda_M. \quad (139)$$

To this aim we start observing that, whenever  $0 \leq k_1 < k_2$  are integers, we have

$$\begin{aligned} |(Du)_{B_{k_2}} - (Du)_{B_{k_1}}| &= \left| \sum_{j=k_1}^{k_2-1} (Du)_{B_{j+1}} - (Du)_{B_j} \right| \\ &\leq \sum_{j=k_1}^{k_2-1} |(Du)_{B_{j+1}} - (Du)_{B_j}| \\ &\leq \sum_{j=k_1}^{k_2-1} \int_{B_{j+1}} |Du - (Du)_{B_j}| dy \\ &\leq \sum_{j=k_1}^{k_2-1} \frac{|B_j|}{|B_{j+1}|} E_j \\ &= \delta_1^{-n} \sum_{j=k_1}^{k_2-1} E_j, \end{aligned} \quad (140)$$

so that by letting  $k_1 = i$  and  $k_2 = h+1$  we obtain

$$a_{h+1} - a_i \leq \delta_1^{-n} \sum_{j=i}^h E_j. \quad (141)$$

By adding up inequalities (137)<sub>3</sub> for  $j \in \{i, \dots, h\}$  and yet adding  $r_i^{1-\alpha} E_i$  to both sides we get

$$\sum_{j=i}^{h+1} r_j^{1-\alpha} E_j \leq r_i^{1-\alpha} E_i + \frac{1}{4} \sum_{j=i}^h r_j^{1-\alpha} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-p+\alpha(p-1)}},$$

so that reabsorbing terms yields

$$\sum_{j=i}^{h+1} r_j^{1-\alpha} E_j \leq 2r_i^{1-\alpha} E_i + 16c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-p+\alpha(p-1)}}.$$

Using then (137)<sub>2</sub> for  $j = i$  and (116) allows to conclude with

$$\sum_{j=i}^{h+1} r_j^{1-\alpha} E_j \leq \frac{\lambda_M \delta_1^n}{20}.$$

The last inequality together (141) gives (recall that  $r_{h+1}^{1-\alpha} \leq r_j^{1-\alpha}$  for  $j \leq h$ )

$$\begin{aligned} r_{h+1}^{1-\alpha} a_{h+1} &\leq r_i^{1-\alpha} a_i + \delta_1^{-n} \sum_{j=i}^h r_j^{1-\alpha} E_j \\ &\leq r_i^{1-\alpha} \int_{B_i} |Du| d\tilde{x} + \delta_1^{-n} \sum_{j=i}^h r_j^{1-\alpha} E_j \\ &\leq \frac{\lambda_M}{100} + \frac{\lambda_M}{20} \leq \frac{\lambda_M}{10}. \end{aligned}$$

Merging the last inequality with (121) and (138), and again using Jensen's inequality, we estimate

$$\begin{aligned} r_{h+1}^{1-\alpha} \int_{B_{h+1}} (|Du| + s) d\tilde{x} &= r_{h+1}^{1-\alpha} \int_{B_{h+1}} (|Du| - |(Du)_{B_{h+1}}|) d\tilde{x} \\ &\quad + r_{h+1}^{1-\alpha} a_{h+1} + r_{h+1}^{1-\alpha} s \\ &\leq r_{h+1}^{1-\alpha} E_{h+1} + r_{h+1}^{1-\alpha} a_{h+1} + R^{1-\alpha} s \\ &\leq \lambda_M, \end{aligned} \tag{142}$$

so that (139) follows. To complete the induction step we shall finally prove (137)<sub>3</sub> with  $j = h + 1$ , that is

$$r_{h+2}^{1-\alpha} E_{h+2} \leq \frac{1}{4} r_{h+1}^{1-\alpha} E_{h+1} + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_h)}{r_h^{n-p+\alpha(p-1)}} \right]. \tag{143}$$

For this we want to apply Lemma 6 in the case  $j = h + 1$ ; let us verify that assumptions are satisfied. Notice that since here we are assuming that  $h \leq i + k - 2$  then  $C_{h+2} \geq \lambda_M/100$ ; moreover, (138) holds and by induction assumption we have

$$r_{j-1}^{1-\alpha} \int_{B_{j-1}} (|Du| + s) d\tilde{x} = r_h^{1-\alpha} \int_{B_h} (|Du| + s) d\tilde{x} \leq \lambda_M.$$

At this point Lemma 6 applies and (143) follows. All in all, using induction (in particular, finite induction when  $k < \infty$ ) this proves

$$\sup_{i \leq j \leq i+k-1} r_j^{1-\alpha} \int_{B_j} (|Du| + s) d\tilde{x} \leq \lambda_M. \quad (144)$$

In particular, this implies (126) when  $k = \infty$ , otherwise we still have to prove that

$$r_{i+k}^{1-\alpha} \int_{B_{i+k}} (|Du| + s) d\tilde{x} \leq \lambda_M. \quad (145)$$

This has been on the other hand implicitly proved before. Indeed, since now (137) holds in the full range  $j \in \{i, \dots, i + k - 1\}$  we repeat the computation from (141) to (142) with  $h = i + k - 1$ , that indeed gives (145). This, together with (144), finally gives (126). The proof is complete.

## 12.7 Proposition 1 and Theorem 8

*Proof of Proposition 1* Let us define  $Q_i = r/2^i$ , for every integer  $i \geq 0$  and  $r = 101|x - y|/100$ . Then observe that

$$\begin{aligned} |(f)_{B(x, Q_i)} - (f)_{B(x, Q_{i+1})}| &\leq \int_{B(x, Q_{i+1})} |f - (f)_{B(x, Q_i)}| d\tilde{x} \\ &\leq 2^n \int_{B(x, Q_i)} |f - (f)_{B(x, Q_i)}| d\tilde{x} \\ &\leq c(n) M_{\alpha, Q_i}^\#(f)(x) Q_i^\alpha \leq c(n) M_{\alpha, R}^\#(f)(x) Q_i^\alpha \end{aligned} \quad (146)$$

and, upon summation, we get

$$\begin{aligned} \sum_{i \geq 0} |(f)_{B(x, Q_i)} - (f)_{B(x, Q_{i+1})}| &\leq c(n) M_{\alpha, R}^\#(f)(x) \sum_{i \geq 0} Q_i^\alpha \\ &\leq \frac{c}{\alpha} M_{\alpha, R}^\#(f)(x) |x - y|^\alpha. \end{aligned} \quad (147)$$

We have used the elementary inequality

$$\sum_{i \geq 0} \frac{1}{2^{i\alpha}} \leq \frac{c}{\alpha}.$$

Notice that this shows that  $\{(f)_{B(x, \varrho_i)}\}$  is a Cauchy sequence and therefore the following limit exists:

$$\lim_{i \rightarrow \infty} (f)_{B(x, \varrho_i)}. \quad (148)$$

Let us now show that the limit

$$\lim_{\varrho \rightarrow 0} (f)_{B(x, \varrho)} \quad (149)$$

exists too, and therefore defines the precise representative of  $f$  and  $x$ , which is indeed denoted by  $f(x)$ . To this aim, take  $0 < \varrho \leq \varrho_0$ . There exists an integer  $i$  such that  $\varrho_{i+1} < \varrho \leq \varrho_i$ ; therefore, as for (146), we have

$$|(f)_{B(x, \varrho_i)} - (f)_{B(x, \varrho)}| \leq c(n) M_{\alpha, R}^{\#}(f)(x) \varrho^{\alpha}.$$

The last inequality, together with the existence of the limit in (148), proves the existence of the limit in (149). Changing in the previous argument  $x$  by  $y$  and choosing this time  $r = |x - y|/100$  gives

$$\sum_{i \geq 0} |(f)_{B(y, \varrho_i)} - (f)_{B(y, \varrho_{i+1})}| \leq \frac{c}{\alpha} M_{\alpha, R}^{\#}(f)(y) |x - y|^{\alpha} \quad (150)$$

and in the same way we can prove that the precise representative of  $f$  can be defined at  $y$ . Notice now that telescoping summation and (147) gives

$$|f(x) - (f)_{B(x, \varrho_0)}| \leq \sum_{i \geq 0} |(f)_{B(x, \varrho_i)} - (f)_{B(x, \varrho_{i+1})}| \leq \frac{c}{\alpha} M_{\alpha, R}^{\#}(f)(x) |x - y|^{\alpha}$$

with a similar estimate for balls centred at  $y$  following by (150). Therefore, using triangle inequality, and recalling the choices of  $r$  made for the points  $x$  and  $y$ , we conclude with

$$\begin{aligned} |f(x) - f(y)| &\leq |(f)_{B(y, 101|x-y|/100)} - (f)_{B(y, |x-y|/100)}| \\ &\quad + \frac{c}{\alpha} \left[ M_{\alpha, R}^{\#}(f)(x) + M_{\alpha, R}^{\#}(f)(y) \right] |x - y|^{\alpha}. \end{aligned}$$

To estimate the first term in the right hand side of the previous inequality we proceed as follows:

$$\begin{aligned}
 & |(f)_{B(x, 101|x-y|/100)} - (f)_{B(y, |x-y|/100)}| \\
 & \leq \int_{B(y, |x-y|/100)} |f - (f)_{B(x, 101|x-y|/100)}| d\tilde{x} \\
 & \leq c \int_{B(x, 101|x-y|/100)} |f - (f)_{B(x, 101|x-y|/100)}| d\tilde{x} \\
 & \leq c M_{\alpha, R}^{\#}(f)(x) |x - y|^{\alpha}.
 \end{aligned}$$

Collecting the inequalities in the last two displays yields (43). Notice that in the last display we have used that  $B(y, |x - y|/100) \subset B(x, 101|x - y|/100)$ . We have also used that

$$\frac{101|x - y|}{100} \leq \frac{404R}{500} \leq R \quad \text{and} \quad B(x, 101|x - y|/100) \subset B_{8R/5}.$$

*Proof of Theorem 8* Applying Proposition 1 in the ball  $B_{R/4}$  and then (42), we get that the estimate

$$\begin{aligned}
 |u(x) - u(y)| & \leq \frac{c}{\alpha} \left[ \mathbf{I}_{p-\alpha(p-1)}^{\mu}(x, R) + \mathbf{I}_{p-\alpha(p-1)}^{\mu}(y, R) \right]^{1/(p-1)} |x - y|^{\alpha} \\
 & \quad + \frac{cR}{\alpha} \int_{B(x, 5R/8)} (|Du| + s) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^{\alpha} \\
 & \quad + \frac{cR}{\alpha} \int_{B(y, 5R/8)} (|Du| + s) d\tilde{x} \cdot \left( \frac{|x - y|}{R} \right)^{\alpha} \quad (151)
 \end{aligned}$$

holds for every choice  $x, y \in B_{R/4}$ . The constant  $c$  appearing in the previous estimate just depends on  $n$ , but is independent of  $\alpha \in [0, 1]$ . In order to estimate the last two integrals, we recall the following Caccioppoli type inequality:

$$\int_B |Du| d\tilde{x} \leq \frac{c}{r} \int_{\gamma B} |u - (u)_{\gamma B}| d\tilde{x} + c \left[ \frac{|\mu|(\gamma B)}{(\gamma r)^{n-1}} \right]^{1/(p-1)} + cs \quad (152)$$

that holds whenever  $\gamma B \subset \Omega$  is a ball with radius  $\gamma r$ , and  $\gamma > 1$ . The constant  $c$  depends only on  $n, p, v, L$  and  $\gamma$ . See [40, Proposition 4.1] for a proof. We apply (152) to  $B(x, 5R/8)$  and  $B(y, 5R/8)$  with the choice  $\gamma = 7/6$  and we use it in (151). Observe that this is possible since

$$x, y \in B_{R/4} \implies \gamma B(x, 5R/8) \cup \gamma B(y, 5R/8) \subset B_R.$$

Then we notice that the following elementary estimate holds whenever  $\sigma \in (0, 1)$ :

$$\begin{aligned} R \left[ \frac{|\mu|(B(x, \sigma R))}{(\sigma R)^{n-1}} \right]^{1/(p-1)} \left( \frac{|x-y|}{R} \right)^\alpha \\ \leq - \frac{1}{\min\{\sigma^{n-1}, \sigma^{(p-1)(1-\alpha)}\} \log \sigma} \left[ \int_{\sigma R}^R \frac{|\mu|(B(x, \varrho))}{\varrho^{n-p+\alpha(p-1)}} \frac{d\varrho}{\varrho} \right]^{1/(p-1)} |x-y|^\alpha \\ \leq c(\sigma) \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) \right]^{1/(p-1)} |x-y|^\alpha \end{aligned}$$

(with a similar one centred at  $y$ ), and, using it with  $\sigma = 35/48$ , we finally conclude with (33).  $\square$

**Remark 7** There is a variant of estimate (33), which is now uniform in the whole range  $\alpha \in [0, 1]$ , and it is the following:

$$\begin{aligned} |u(x) - u(y)| \leq \frac{c}{\alpha} \left[ \mathbf{I}_{p-\alpha(p-1)}^\mu(x, R) + \mathbf{I}_{p-\alpha(p-1)}^\mu(y, R) \right]^{1/(p-1)} |x-y|^\alpha \\ + \frac{c}{\alpha} \int_{B_R} (|u| + Rs) d\tilde{x} \cdot \left( \frac{|x-y|}{R} \right)^\alpha \end{aligned} \quad (153)$$

The constant  $c$  appearing in the above estimate is this time independent of  $\alpha$ , and only exhibits a dependence on  $n, p, \nu, L$ . Estimate (153) follows exactly as estimate (33), but tracking the dependence on the constants on  $\alpha$  starting from (151). Estimate (153) describes in a precise way the blow-up rate of the constant  $c$  in (33) when  $\alpha \rightarrow 0$ . More precisely, the coefficient  $1/\alpha$  appears to be related to the loss of  $L^\infty$ -estimates towards BMO-ones already discussed after Proposition 1. Moreover, it is the same blow-up rate appearing when fractional estimates get lost as the fractional differentiability parameter disappears. See for instance [52].

### 13 Proof of Theorem 12

This theorem has been proved in [40] in a slightly different form and with a different proof; Theorem 12 is essential for the proof of Theorem 9. We consider the sequence of shrinking balls introduced in (81) with the corresponding solutions  $v_j \in u + W_0^{1,p}(B_i)$  defined in (112); the number  $\delta_1 \in (0, 1/4)$  is this time chosen as

$$\delta_1 := \left( \frac{\varepsilon}{4c_d} \right)^{1/(\beta-\tilde{\alpha})}, \quad \beta := \frac{\alpha_M + \tilde{\alpha}}{2}, \quad (154)$$

while we take  $r := R$ . Here we take  $\varepsilon \in (0, 1/2)$  to be a fixed number. Finally, we recall the definition of the excess quantities  $\{E_j\}$  in (117). Before going on let us



record an identity of later use, that is

$$r_j^{-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} = \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}. \quad (155)$$

We are now going to use estimate (105) as we have done for Lemma 4. In particular, in light of Remark 5, with  $\tilde{\alpha} < \alpha_M$ , we can choose in (105) a number  $\beta$  such that  $\tilde{\alpha} < \beta < \alpha_M$  and we take the one in (154). Therefore, recalling that  $r_{j+1}^{-\alpha} = \delta_1^{-\alpha} r_j^{-\alpha}$ , we have

$$\begin{aligned} r_{j+1}^{-\alpha} E_{j+1} &\leq 2r_{j+1}^{-\alpha} \int_{B_{j+1}} |Dv_j - (Dv_j)_{B_{j+1}}| d\tilde{x} \\ &\quad + 2\delta_1^{-n} r_{j+1}^{-\alpha} \int_{B_j} |Dv_j - Du| d\tilde{x} \\ &\leq 2c_d \delta_1^{\beta-\alpha} r_j^{-\alpha} \int_{B_j} |Dv_j - (Dv_j)_{B_j}| d\tilde{x} \\ &\quad + 2c\delta_1^{-n-\alpha} r_j^{-\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \\ &\leq 4c_d \delta_1^{\beta-\tilde{\alpha}} r_j^{-\alpha} E_j + 8c\delta_1^{-n-\tilde{\alpha}} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}. \end{aligned}$$

By recalling (154) we conclude with

$$r_{j+1}^{-\alpha} E_{j+1} \leq \varepsilon r_j^{-\alpha} E_j + c \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \quad (156)$$

where  $c \equiv c(n, p, \nu, L, \tilde{\alpha})$  as a consequence of the fact that the choice in (154) determines a number  $\delta_1$  which again depends on  $n, p, \nu, L, \tilde{\alpha}$  and  $\varepsilon$  via the dependence of the quantities  $\beta$  and  $c_d$ . Observe that the resulting estimate is independent of  $\alpha$  as long as this belongs to  $[0, \tilde{\alpha}]$ . From (156) it follows that

$$\sup_{j \geq 1} r_j^{-\alpha} E_j \leq \varepsilon \sup_{j \geq 1} r_j^{-\alpha} E_j + \varepsilon R^{-\alpha} E_0 + c(\varepsilon) \sup_{j \geq 0} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}. \quad (157)$$

Taking for instance  $\varepsilon = 1/4$ , reabsorbing terms and yet adding  $R^{-\alpha} E_0$  to both sides of the resulting inequality, we conclude with

$$\begin{aligned} \sup_{j \geq 0} r_j^{-\alpha} E_j &\leq c R^{-\alpha} \int_{B_R} |Du - (Du)_{B_R}| d\tilde{x} \\ &\quad + c [M_{1-\alpha(p-1), R}(\mu)(x)]^{1/(p-1)} =: cS. \end{aligned} \quad (158)$$

This last estimate implies the statement of the theorem, that is

$$M_{\alpha, R}^{\#}(Du)(x) \leq cS, \quad (159)$$

where  $c \equiv c(n, p, \nu, L, \tilde{\alpha})$  and  $S$  in fact denotes the quantity appearing in the right hand side of (158). Indeed, let us consider a positive radius  $\varrho \leq R$ , and determine  $j \geq 0$  such that  $r_{j+1} < \varrho \leq r_j$ ; then using also (98) we have

$$\begin{aligned} \varrho^{-\alpha} \int_{B_{\varrho}} |Du - (Du)_{B_{\varrho}}| d\tilde{x} &\leq 2\varrho^{-\alpha} \int_{B_{\varrho}} |Du - (Du)_{B_j}| d\tilde{x} \\ &\leq 2 \left( \frac{\varrho}{r_j} \right)^{-n-\alpha} r_j^{-\alpha} \int_{B_j} |Du - (Du)_{B_j}| d\tilde{x} \\ &\leq 2\delta_1^{-n-\tilde{\alpha}} r_j^{-\alpha} \int_{B_j} |Du - (Du)_{B_j}| d\tilde{x} \leq cS \end{aligned}$$

and (159) follows keeping in mind the dependence of  $\delta_1$  on  $n, p, \nu, L, \tilde{\alpha}$ . The proof is complete.

### 13.1 A VMO-type result

A modification to the proof of Theorem 12 allows to get a corollary that is interesting in itself and that will be useful in order to obtain subsequent results.

**Theorem 19** (VMO gradient regularity) *Let  $u \in W^{1,p}(\Omega)$  be a weak solution to the Eq. (13) under the assumptions (16). Assume that*

$$S_1(x) := \sup_{B(x, \varrho) \subset \Omega} \int_{B(x, \varrho)} |Du| d\tilde{x} < \infty \quad (160)$$

and

$$\lim_{R \rightarrow 0} \frac{|\mu|(B(x, R))}{R^{n-1}} = 0 \quad (161)$$

hold for a certain  $x \in \Omega$ , then

$$\lim_{R \rightarrow 0} \int_{B(x, R)} |Du - (Du)_{B(x, R)}| d\tilde{x} = 0. \quad (162)$$

Moreover, if  $Du$  is locally bounded and if the limit in (161) is uniform with respect to  $x$  then  $Du$  is (locally) VMO in  $\Omega$ .

*Proof* In order to prove (162) we show that, for every choice of  $\tilde{\varepsilon}$ , there exists  $\tilde{R} > 0$ , depending on  $n, p, \nu, L, \tilde{\varepsilon}, \mu$  and the point  $x$ , such that

$$\varrho \leq \tilde{R} \implies E(Du, B(x, \varrho)) < \tilde{\varepsilon}. \quad (163)$$

We revisit the proof of Theorem 12, where we consider the case  $\alpha = \tilde{\alpha} = 0$ . Assume that  $B(x, R) \subset \Omega$ . Estimate (157) then gives

$$\sup_{j \geq 1} E_j \leq 2\varepsilon E(Du, B(x, R)) + c(\varepsilon) [M_{1,R}(\mu)(x)]^{1/(p-1)},$$

which is valid as soon as  $\delta_1$  and  $\varepsilon$  are linked as in (154), and  $\varepsilon \in (0, 1/2)$ . In particular, we also have

$$\sup_{j \geq 1} E_j \leq 4\varepsilon S_1 + c(\varepsilon) [M_{1,R}(\mu)(x)]^{1/(p-1)}, \quad (164)$$

where we have used the abbreviation  $S_1 \equiv S_1(x)$ . Now, let us notice how the previous argument implies that

$$\sup_{\varrho \leq \delta_1 R} E(Du, B(x, \varrho)) \leq 4\varepsilon S_1 + c(\varepsilon) [M_{1,R}(\mu)(x)]^{1/(p-1)} \quad (165)$$

holds for every  $\varrho \leq \delta_1 R = r_1$  and  $\delta_1$  is again chosen as in (154). Indeed, consider a number  $\varrho \leq \delta_1 R$  with  $\tilde{\alpha} = 0$ ; then there exists an integer  $k \geq 1$  such that  $r_{k+1} < \varrho \leq r_k$ . This means that  $\varrho = \delta_1^k R'$  for some  $R' \in (\delta_1 R, R]$ . Now, in the proof of Theorem 12 replace  $R$  by  $R'$  and we gain a new chain of cylinders  $B_j$  for which (164) obviously holds; indeed, notice that  $M_{1,R'}(\mu)(x) \leq M_{1,R}(\mu)(x)$  since  $R' \leq R$ . With this new definition we have  $E(Du, B_\varrho) = E_k$  and therefore (165) is nothing but (164) with this new choice of the radii  $\{r_j\} \equiv \{\delta_1^j R'\}$ . All in all we have proved that (165) holds with  $\delta_1$  depending only on  $n, p, \nu, L$  and  $\varepsilon$ . Let us recall that estimate (165) holds with any choice of the initial radius  $R$ , which is a free parameter that is going to be chosen in the next lines; notice indeed that all the constants determined up to now are independent of the starting radius  $R$ . Now, first choose  $\varepsilon \equiv \varepsilon(n, p, \nu, L, \tilde{\varepsilon})$  small enough in order to have

$$4\varepsilon S_1 \leq \frac{\tilde{\varepsilon}}{2}. \quad (166)$$

This determines the constant  $c(\varepsilon)$  as a function of  $n, p, \nu, L, \tilde{\varepsilon}$ ; notice that this is possible by the first assumption in (161), while using the second we can determine a radius  $R_1$ , depending essentially on  $n, p, \nu, L, \tilde{\varepsilon}$  and the point  $x$  (via the rate of convergence in (161)), such that

$$R \leq R_1 \implies c(\varepsilon) [M_{1,R}(\mu)(x)]^{1/(p-1)} \leq \frac{\tilde{\varepsilon}}{2}. \quad (167)$$

Combining the content of the last two displays with the one of (165) (where we now take with  $R \equiv R_1$ ), and recalling the definition of the excess functional  $E(Du, B_R)$ , yields (163) with the choice  $\tilde{R} = \delta_1 R_1$ . As for the local VMO-regularity of  $Du$  we recall that this means that, for every  $\tilde{\varepsilon}$ , the choice of  $\tilde{R}$  in (163) can be done independently of the point  $x$  as long as  $B_{\tilde{R}}$  varies in a fixed compact subset  $\Omega' \Subset \Omega$ . We can now check this by showing that all the choices of the constants above can be done independently of the point  $x$ , but depending only on the fixed subset  $\Omega'$ . Indeed (166) can be replaced by

$$4\varepsilon S_1 \leq 4\varepsilon \|Du\|_{L^\infty(\Omega')} \leq \frac{\tilde{\varepsilon}}{2}. \quad (168)$$

The choice of  $R_1$  in (167) can be made uniform as well by (161), which is now uniform with respect to  $x$ , and the proof of the VMO-regularity of  $Du$  is complete, too.

## 14 Proof of Theorem 15

We first prove the theorem in the case we have an energy solution  $u \in W^{1,p}(\Omega)$ ; this will be done through Steps 14.1–14.3. Then, in the final Step 14.4, we describe the approximation argument to prove the theorem for any SOLA  $u \in W_0^{1,p-1}(\Omega)$  to problem (55). We indeed prefer a separate treatment of the two cases since in this second part we shall also present the general arguments to prove that the other potential estimates presented in this paper for energy solutions actually hold for SOLA too.

### 14.1 Proof in the case of energy solutions: beginning

We fix a ball  $B_R \equiv B(x, R) \subset \Omega$  and then consider in the rest of the proof balls which are concentric to it. We notice that by assumption (60) the quantity

$$\lambda_M := c [\mathbf{I}_1^\mu(x, R)]^{1/(p-1)} + c \int_{B_R} (|Du| + s) d\tilde{x} \quad (169)$$

is finite, where  $c \equiv c(n, p, \nu, L)$  is the constant appearing in (40). We shall then prove that, for every  $\varepsilon > 0$ , there exists a radius  $r_\varepsilon \leq R$ , in general depending only on  $n, p, \nu, L$  and the point  $x$ , such that

$$|(Du)_{B(x,\varrho)} - (Du)_{B(x,\tau)}| \leq \lambda_M \varepsilon \quad \text{holds whenever } 0 < \tau < \varrho \leq r_\varepsilon. \quad (170)$$

This will finally prove (61). Notice that with the above definitions, by (40) we have

$$\oint_{B(x, \varrho)} (|Du| + s) d\tilde{x} \leq \lambda_M \quad \text{for every } \varrho \leq R \quad (171)$$

while, and this time by Theorem 19, we have that

$$\lim_{\varrho \rightarrow 0} E(Du, B(x, \varrho)) = 0. \quad (172)$$

Notice that, in order to verify the assumptions in (161) we have implicitly used the fact that  $\mathbf{I}_1^\mu(x, R) < \infty$  implies that

$$\lim_{\varrho \rightarrow 0} \mathbf{I}_1^\mu(x, \varrho) = 0 \quad (173)$$

and this, thanks to an argument similar to the one on (115), also implies

$$\lim_{\varrho \rightarrow 0} \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1}} = 0 \quad (174)$$

which is in fact the second part of the main assumption in Theorem 161.

## 14.2 Setting of the constants and a sequence

We revisit the arguments of the proof of Theorem 10, while in the following the number  $\varepsilon$  is fixed as the one introduced in (170). As usual the constants  $c_l, c_h, c_d$  have been defined in (100), (101) and (105), respectively;  $c_1$  is from Lemma 2 (where we take  $q = 1$ ) and  $\beta \equiv \beta(n, p, v, L) \in (0, 1)$  appears in (101) and (105). We this time fix  $\delta_1 \equiv \delta_1(n, p, v, L, \varepsilon)$  as

$$\delta_1 := \left( \frac{\varepsilon^2}{10^{8n} c_l^2 c_d c_h} \right)^{1/\beta}. \quad (175)$$

With such a choice of  $\delta_1$  we determine the constant  $c_2$  from Lemma 3 as follows:

$$c_2 := c_2(n, p, v, L, \delta_1, A) \quad \text{with} \quad A := \frac{10^{4n} c_l}{\varepsilon}. \quad (176)$$

Notice that all the constants appearing up to now show a global dependence on the constants  $n, p, v, L$  and  $\varepsilon$ . Now we can take a radius  $R_l \equiv R_l(n, p, v, L, \varepsilon, x) \leq R$  such that the following two conditions are satisfied:

$$\left[ \int_0^{2R_l} \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1}} \frac{d\varrho}{\varrho} \right]^{1/(p-1)} \leq \frac{\delta_1^{5n} \lambda_M \varepsilon}{10^8 c_1 c_2 c_d} \quad (177)$$

and

$$\sup_{0 < \varrho \leq R_l} E(Du, B(x, \varrho)) \leq \frac{\delta_1^{5n} \lambda_M \varepsilon}{10^8}. \quad (178)$$

Notice that we can satisfy (177) by (173), while the choice in (178) is possible due to (172). Finally, for integers  $j \geq 0$  we consider the balls as in (81) where this time

$$B_j := B(x, r_j), \quad r_j = \delta_1^j r, \quad r := R_l$$

and the related comparison solutions  $v_j \in u + W_0^{1,p}(B_j)$  in (112). Accordingly, the numbers  $\{E_i\}$  are again defined as in (117). Notice also that (177) and a computation totally similar to the one in (115) (where we take  $\alpha = 1$ ) give

$$\left[ \sum_{j=0}^{\infty} \frac{|\mu|(\bar{B}_j)}{r_j^{n-1}} \right]^{1/(p-1)} \leq \frac{\delta_1^{4n} \lambda_M \varepsilon}{10^8 c_1 c_2 c_d}. \quad (179)$$

In Step 14.3 below, we will prove that

$$|(Du)_{B_h} - (Du)_{B_k}| \leq \frac{\lambda_M \varepsilon}{12} \quad \text{holds whenever } 2 \leq k \leq h. \quad (180)$$

Let us assume for a moment that (180) and let us finish the proof of (170) with the choice

$$r_\varepsilon := \delta_1^2 R_l. \quad (181)$$

Indeed, let us fix  $0 < \tau < \varrho \leq r_\varepsilon$ . This means that there exist two integers  $k$  and  $h$ , such that  $2 \leq k \leq h$ ,

$$\delta_1^{k+1} R_l < \varrho \leq \delta_1^k R_l \quad \text{and} \quad \delta_1^{h+1} R_l < \tau \leq \delta_1^h R_l \quad (182)$$

hold. Applying (178), and taking (182) into account, we get

$$\begin{aligned} |(Du)_{B(x, \varrho)} - (Du)_{B_{k+1}}| &\leq \int_{B_{k+1}} |Du - (Du)_{B(x, \varrho)}| d\tilde{x} \\ &\leq \frac{|B(x, \varrho)|}{|B_{k+1}|} \int_{B(x, \varrho)} |Du - (Du)_{B(x, \varrho)}| d\tilde{x} \\ &\leq \delta_1^{-n} E(Du, B(x, \varrho)) \leq \frac{\lambda_M \varepsilon}{10}. \end{aligned}$$

Similarly, we also gain

$$|(Du)_{B(x,\tau)} - (Du)_{B_{h+1}}| \leq \frac{\lambda_M \varepsilon}{10}.$$

Using the inequalities in the last two displays together with (180) and triangle inequality establishes (170). It remains to prove (180) and this will be done in the next step.

### 14.3 Non-degenerate iteration chains and proof of (180)

We set, for  $j \geq 1$

$$C_j := \int_{B_j} |Du| d\tilde{x}$$

and then, as in (122), we define

$$\mathcal{L} := \left\{ j \in \mathbb{N} : C_j \leq \frac{\lambda_M \varepsilon}{100} \right\}. \quad (183)$$

Accordingly, we also define the chains  $C_i^k$  and  $C_i^\infty$  as in (123) and (124), respectively. The difference here is the appearance of the parameter  $\varepsilon$  in the new definition of the set  $\mathcal{L}$  in (183); this allows to interpret the sets  $C_i^k$  as maximal chains along which the equation becomes nondegenerate. In this case it will be useful to consider the following number:

$$j_m := \min \mathcal{L}. \quad (184)$$

Needless to say, in case  $\mathcal{L}$  is empty we have  $j_m = \infty$  and this is actually a favourable case since the problem never becomes degenerate. The following lemma is similar to Lemma 6:

**Lemma 7** *Let  $j \geq 1$  be an integer such that*

$$C_{j+1} \geq \frac{\lambda_M \varepsilon}{100}$$

*holds; then we have*

$$E_{j+1} \leq \frac{\varepsilon}{4} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right], \quad (185)$$

*where  $c_2$  depends only on  $n, p, v, L, \varepsilon$  and has been defined in (176).*

*Proof* As for Lemma 6 we estimate

$$\begin{aligned}
 \int_{B_{j+1}} |Dv_{j-1}| d\tilde{x} &\geq C_{j+1} - \int_{B_{j+1}} |Du - Dv_{j-1}| d\tilde{x} \\
 &\geq \frac{\lambda_M \varepsilon}{100} - \int_{B_{j+1}} |Du - Dv_{j-1}| d\tilde{x} \\
 &\geq \frac{\lambda_M \varepsilon}{100} - \delta_1^{-2n} \int_{B_{j-1}} |Du - Dv_{j-1}| d\tilde{x} \\
 &\geq \frac{\lambda_M \varepsilon}{100} - \delta_1^{-2n} c_1 \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]^{1/(p-1)} \\
 &\geq \frac{\lambda_M \varepsilon}{200}
 \end{aligned} \tag{186}$$

where we used (76) and (179). On the other hand, using triangle inequality, (171), (179) and yet Lemma 2, we have

$$\int_{B_{j-1}} (|Dv_{j-1}| + s) d\tilde{x} \leq \int_{B_{j-1}} (|Du| + s) d\tilde{x} + c_1 \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right]^{1/(p-1)} \leq 2\lambda_M.$$

Applying estimate (100) with  $v \equiv v_{j-1}$  then yields

$$(|Dv_{j-1}| + s) \leq 2^{n+1} c_l \lambda_M \quad \text{in } B_{j-1}/2. \tag{187}$$

In view of (186)–(187) we are in position to apply Proposition 2 with the choices  $v \equiv v_{j-1}$ ,  $B \equiv B_{j-1}$ ,  $\sigma = \delta_1$ ,  $k = 1$ ,  $\lambda \equiv \lambda_M$  and  $\Gamma = 10^{3n} c_l / \varepsilon$ ; observe that this is again possible by the choice of  $\delta_1$  made in (175). This yields

$$\frac{\lambda_M \varepsilon}{10^{4n} c_l} \leq |Dv_{j-1}| \quad \text{in } B_j$$

The inequality in the above display and (187) summarise in the following line:

$$\frac{\lambda}{A} \leq |Dv_{j-1}| \leq A\lambda \quad \text{in } B_j, \quad A = \frac{10^{4n} c_l}{\varepsilon}$$

and therefore we are in position to apply Lemma 3 thereby obtaining

$$\int_{B_j} |Du - Dv_j| d\tilde{x} \leq c_2 \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right],$$



where the constant  $c_2$  has been fixed in (176). With the last comparison estimate we can conclude with (185) as in (136).  $\square$

We now proceed with the proof of (180), obviously assuming  $k < h$ . We then analyse three different cases.

*Case 1:*  $k < h \leq j_m$ , where we recall that the number  $j_m$  has been defined in (184). By the very definition of  $j_m$  it follows that  $C_{j+1} \geq \lambda_M \varepsilon / 100$  holds for every  $j \in \{k-1, \dots, h-2\}$  and therefore we apply Lemma 7 to get that

$$E_{j+1} \leq \frac{1}{4} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \left[ \frac{|\mu|(B_{j-1})}{r_{j-1}^{n-1}} \right] \quad (188)$$

holds, again for every  $j \in \{k-1, \dots, h-2\}$ . Summing up yields

$$\sum_{j=k}^{h-1} E_j \leq \frac{1}{4} \sum_{j=k-1}^{h-2} E_j + 8c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-1}}$$

so that

$$\sum_{j=k}^{h-1} E_j \leq \frac{1}{3} E_{k-1} + 16c_2 c_d \delta_1^{-n} \lambda_M^{2-p} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-1}}.$$

Finally, recalling (178) and (179), we conclude with

$$\sum_{j=k}^{h-1} E_j \leq \frac{\delta_1^{2n} \lambda_M \varepsilon}{50}.$$

On the other hand, taking  $k_1 = k$  and  $k_2 = h$  in (140) and using the inequality in the last display we can estimate

$$|(Du)_{B_h} - (Du)_{B_k}| \leq \delta_1^{-n} \sum_{j=k}^{h-1} E_j \leq \lambda_M \varepsilon \quad (189)$$

and (180) follows.

*Case 2:*  $j_m \leq k < h$ . Here (180) follows as a consequence of the inequalities

$$|(Du)_{B_h}| \leq \frac{\lambda_M \varepsilon}{25} \quad \text{and} \quad |(Du)_{B_k}| \leq \frac{\lambda_M \varepsilon}{25}. \quad (190)$$

If  $h \in \mathcal{L}$ , the first inequality in (190) follows immediately from the definition of  $\mathcal{L}$ ; we can therefore assume  $h \notin \mathcal{L}$ . Then, as  $h > j_m$ , it is possible to consider a nondegenerate iteration chain  $\mathcal{C}_{i_h}^{m_h}$  with  $m_h > 0$ , such that  $h \in \mathcal{C}_{i_h}^{m_h}$ ; notice that  $h > i_h$  as  $h \notin \mathcal{L} \ni i_h$ , since by definition the first index in a chain does belong to  $\mathcal{L}$ . We are again in position to apply Lemma 7 for  $j \in \{i_h, \dots, i_h + m_h - 1\}$ , so that (188)

holds for the corresponding indexes. Summing up, proceeding as after (188), and yet summing  $E_{i_h}$  to both sides of the resulting inequality, we arrive at

$$\sum_{j=i_h}^{i_h+m_h} E_j \leq 2E_{i_h} + 8c_2c_d\delta_1^{-n}\lambda_M^{2-p} \sum_{j=0}^{\infty} \frac{|\mu|(B_j)}{r_j^{n-1}} \leq \frac{\delta_1^{2n}\lambda_M\varepsilon}{50}.$$

We have again used (178) and (179). Proceeding as in (140) (taking  $k_1 = i_h$  and  $k_2 = h$ ), and using the definition of  $C_{i_h}^{m_h}$  to estimate

$$|(Du)_{B_{i_h}}| \leq C_{i_h} \leq \frac{\lambda_M\varepsilon}{100}$$

we have

$$\begin{aligned} |(Du)_{B_h}| &\leq |(Du)_{B_h} - (Du)_{B_{i_h}}| + |(Du)_{B_{i_h}}| \\ &\leq \delta_1^{-n} \sum_{i=i_h}^{h-1} E_i + \frac{\lambda_M\varepsilon}{100} \\ &\leq \delta_1^{-n} \sum_{i=i_h}^{i_h+m_h} E_i + \frac{\lambda_M\varepsilon}{100} \leq \frac{\lambda_M\varepsilon}{25}, \end{aligned}$$

that is the first inequality in (190). The proof of the second inequality in (190) is completely similar; we just observe that we can assume that  $k > j_m$  otherwise the inequality itself is trivial as  $k \in \mathcal{L}$ .

*Case 3:*  $k < j_m < h$ . This can be actually treated by a combination of the first two cases. It is indeed sufficient to prove that the inequalities in display (190) still hold. The first inequality in (190) follows exactly as in Case 2. As for the second estimate in (190), let us remark that, as  $j_m \in \mathcal{L}$ , we have that

$$|(Du)_{B_{j_m}}| \leq C_{j_m} \leq \frac{\lambda_M\varepsilon}{50}.$$

On the other hand, we can repeat the argument of Case 1, with  $h$  replaced by  $j_m$ , iterating from  $h$  to  $j_m$  and thereby obtaining

$$|(Du)_{B_{j_m}} - (Du)_{B_k}| \leq \frac{\lambda_M\varepsilon}{50}$$

as in (189). Finally, the inequalities in the last two displays and yet triangle inequality give

$$|(Du)_{B_k}| \leq |(Du)_{B_k} - (Du)_{B_{j_m}}| + |(Du)_{B_{j_m}}| \leq \frac{\lambda_M\varepsilon}{50} + \frac{\lambda_M\varepsilon}{100} \leq \frac{\lambda_M\varepsilon}{25}$$

and therefore the second inequality in (190) follows. The proof of Theorem 15 is therefore complete in the case of an energy solution  $u \in W^{1,p}(\Omega)$ .

## 14.4 Very weak solutions

In this section we are going to prove Theorem 15 for SOLA. We shall also obtain the extensions of Theorems 10 and 19 to SOLA, since these are indeed needed as preliminary results. The method we propose here is of course based on an approximation argument, but, rather than passing to the limits in the final estimates, we will pass to the limits “in the proofs”. Strictly speaking, we shall pass to the limits in some estimates from the proofs. The same methods developed for energy solutions will then work for SOLA. We start with a preliminary result. Let us consider a ball  $B(x, R) \subset \Omega$ , and a sequence of shrinking balls  $\{B_j\}$ , concentric to  $B(x, R)$ :

$$\cdots B_{j+1} \subset B_j \subset B_{j-1} \cdots B_0 \subset B(x, R). \quad (191)$$

Then we have the following:

**Lemma 8** *Let  $u$  be a SOLA to (17) under the assumptions (18), and let  $\{B_j\}$  be the sequence of balls considered in (191). There exists a sequence of functions  $\{v_j\}$  such that for every  $j \geq 0$  it holds*

- $v_j \in W_{\text{loc}}^{1,p}(B_j) \cap u + W_0^{1,p-1}(B_j)$
- $v_j$  is a local energy solution in  $B_j$  i.e.  $\operatorname{div} a(x, Dv_j) = 0$  in  $B_j$
- There exist constants  $c_1, \bar{c}_1 \equiv c_1, \bar{c}_1(n, p, v)$  such that

$$\int_{B_j} |Du - Dv_j|^q d\tilde{x} \leq c_1 \left[ \frac{|\mu|(\bar{B}_j)}{r_j^{n-1}} \right]^{q/(p-1)} \quad (192)$$

and

$$\int_{B_j} |u - v_j|^q d\tilde{x} \leq \bar{c}_1 \left[ \frac{|\mu|(\bar{B}_j)}{r_j^{n-p}} \right]^{q/(p-1)} \quad (193)$$

hold whenever  $q$  is as in (78)

- There exists a constant  $c \equiv c(n, p, v)$  such that

$$\int_{B_j} \frac{|V(Du) - V(Dv_j)|^2}{(h + |u - v_j|)^\xi} d\tilde{x} \leq c \frac{h^{1-\xi}}{\xi - 1} |\mu|(\bar{B}_j) \quad (194)$$

holds whenever  $h > 0$  and  $\xi > 1$ , where  $c \equiv c(n, p, v) \geq 1$ . The function  $V(\cdot)$  has been defined in (71).

*Proof* The proof goes via approximation and we confine ourselves to explain the arguments for the inequality in display (192), those for the other ones being completely analogous. We preliminary recall that by the results in [8, 55]—see also Theorem 14—any SOLA  $u$  to (55) belongs to  $W_{\text{loc}}^{1,q}(\Omega)$  for the numbers  $q$  in the range described in

(76). With  $u$  being now a fixed SOLA to (17), we consider the associated sequence  $\{u_k\} \subset W_{\text{loc}}^{1,p}(\Omega)$  from Definition 8 and note that using a standard diagonal we may assume, up to passing to not relabelled subsequences, that  $Du_k \rightarrow Du$  a.e.. With  $B_j$  being fixed we now build the function  $v_j$  as required in the statement of the lemma. We start defining  $w_k \in u_k + W_0^{1,p}(B_j)$  as the unique solution to the Dirichlet problem

$$\begin{cases} \operatorname{div} a(x, Dw_k) = 0 & \text{in } B_j \\ w_k = u_k & \text{on } \partial B_j. \end{cases} \quad (195)$$

Applying Lemma 2 in this context gives

$$\int_{B_j} |Du_k - Dw_k|^q d\tilde{x} \leq c_1 \left[ \frac{|\mu_k|(\bar{B}_j)}{r_j^{n-1}} \right]^{q/(p-1)}. \quad (196)$$

The next step is now letting  $k \rightarrow \infty$  in the previous inequality, and for this we need to recall a few regularity properties of the functions  $w_k$ . An important point is that the way the functions  $w_k$  have been defined gives that they belong to the Sobolev space  $W^{1,p}(B_j)$  and this allows to apply the regularity theory available for such solutions. By (196) and the very definition of the functions  $u_k$  (they are converging to  $u$  in  $W_0^{1,p-1}(\Omega)$ ) it follows that the sequence  $\{w_k\}$  is bounded in  $W^{1,p-1}(B_j)$ ; in turn this fact and the bound in (100) allows to get a uniform bound of the type  $\|Dw_k\|_{L^\infty(\gamma B_j)} \leq c(\gamma)$ . In the same way we deduce, by using (101), that the sequence of maps  $\{Dw_k\}$  is equicontinuous in  $\gamma B_j$ . We can therefore invoke Ascoli–Arzelà’s theorem that, together with a standard diagonal argument (recall that here  $\gamma < 1$  is arbitrary), gives that there exists a function  $w \in W_{\text{loc}}^{1,\infty}(B_j) \cap u + W_0^{p-1}(B_j)$  such that, up to a not relabelled subsequence, we have  $w_k \rightarrow w$  and  $Dw_k \rightarrow Dw$  locally uniformly in  $B_j$ . Now, the locally uniform convergence of the gradients  $Dw_k$  allows to pass to the limits in the equations (195) and to conclude that  $\operatorname{div} a(x, Dw) = 0$ . Fatou’s lemma now gives

$$\begin{aligned} \int_{B_j} |Du - Dw|^q d\tilde{x} &\leq \liminf_{k \rightarrow \infty} \int_{B_j} |Du_k - Dw_k|^q d\tilde{x} \\ &\leq \limsup_{k \rightarrow \infty} c_1 \left[ \frac{|\mu_k|(B_j)}{r_j^{n-1}} \right]^{q/(p-1)} \leq c_1 \left[ \frac{|\mu|(\bar{B}_j)}{r_j^{n-1}} \right]^{q/(p-1)}, \end{aligned}$$

and in the last estimate we have used a standard property of the weak\*-convergence of measures. A similar reasoning can be worked out with respect to (193) and (194). All in all we have proved the lemma with the choice  $v_j \equiv w$ .  $\square$

With the previous lemma at our disposal the proofs of Theorems 10, 19 and 15 for SOLA proceed verbatim as in the case of standard energy solutions, with the sequence  $v_j$  replacing everywhere the one in (112) whenever this is used. Indeed all the remaining facts used about the functions  $v_j$  follow by Lemma 8. In particular, all

the facts reported in Sects. 11 and 12.1 are still valid since the functions  $v_j$  belong to  $W_{\text{loc}}^{1,p}(B_j) \cap W^{1,p-1}(B_j)$  and therefore standard regularity theory applies to them. The only difference is that the quantities  $|\mu|(\bar{B}_j)$  must replace everywhere the analogous  $|\mu|(B_j)$ . This is clearly not a problem since the basic inequalities the quantities  $|\mu|(\bar{B}_j)$  have to satisfy are implied by (116) and (179), which are already involving  $|\mu|(\bar{B}_j)$ . Needless to say, whenever we are considering vanishing limits as in (161) the same thing hold when considering closed balls.

### 15 Theorem 3 as a corollary

The assertion of Theorem 3 can be now obtained as a corollary of the results presented up to now, and, by the approach in the previous section, directly for SOLA. Indeed, estimate (24) follows directly by estimate (40) and the fact that the finiteness of  $\mathbf{I}_1^\mu(x, R)$  implies that  $x$  is a Lebesgue point of  $Du$ . This is ensured by Theorem 15, see also Remark 1. It remains to prove that (25) implies the continuity of  $Du$ . For this we shall essentially revisit the proof of Theorem 15 given in the previous section and we shall see that the Cauchy sequence information in (170) holds locally uniformly in  $x$ . This means that the radius  $r_\varepsilon$  can be chosen independently of the point  $x$  as long as this varies in a fixed open subset  $\Omega' \Subset \Omega$ . The continuity of  $Du$  then follows by showing that it can be obtained as the locally uniform limit of the net of continuous maps defined by

$$x \rightarrow (Du)_{B(x,\varrho)} \quad \text{for } 0 < \varrho \leq d := \text{dist}(\Omega', \partial\Omega), \quad (197)$$

which are obviously continuous. Note that we already know that the above maps converge pointwise to  $Du$  (that is, they converges to its precise representative at every point) as established in Theorem 15. Now, a standard covering argument and the assumption (25) imply that  $Du$  is locally bounded in  $\Omega$ . Therefore, selecting another open subset  $\Omega''$  such that  $\Omega' \Subset \Omega'' \Subset \Omega$ , we set

$$\lambda_M := c \|\mathbf{I}_1^\mu(\cdot, R)\|_{L^\infty(\Omega'')}^{1/(p-1)} + c \| |Du| + s \|_{L^\infty(\Omega'')} \quad (198)$$

for some  $R > 0$ , where  $c$  is the constant appearing in (40). This number is going to replace the analogous one picked in (169), making it independent of  $x$ . Then, we observe that assuming (25) allows to satisfy condition (174), uniformly in  $x$  (with an argument for instance based on computations as in (115)) and this implies that (172) holds again locally uniformly in  $x$  via Theorem 19. This, and the assumption (25), imply that we can replace (177) and (178) by the stronger conditions

$$\sup_{x \in \Omega'} \left[ \int_0^{2R_l} \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1}} \frac{d\varrho}{\varrho} \right]^{1/(p-1)} \leq \frac{\delta_1^{5n} \lambda_M \varepsilon}{10^8 c_1 c_2 c_d}$$

and

$$\sup_{x \in \Omega'} \sup_{0 < \varrho \leq R_l} E(Du, B(x, \varrho)) \leq \frac{\delta_1^{5n} \lambda_M \varepsilon}{10^8}$$

respectively, where now in addition we take  $R_l < \text{dist}(\Omega', \partial\Omega'')$  and  $\lambda_M$  is defined as in (198). With the inequalities of the last two displays being now in force it is easy to check that all the inequalities after (179) become uniform with respect to the point  $x$  considered. In particular, the numbers  $\delta_1$  and  $R_l$  that determine  $r_\varepsilon$  via (181), that is the radius for which (170) holds, are independent of  $x$ . This means that the maps defined in (197) converge uniformly to  $Du$  and the continuity of  $Du$  therefore follows.

## 16 Proof of Theorem 9

Theorem 9 has been again proved in a slightly different form in [40]. The proof of Theorem 9 goes essentially in two stages: we first obtain a pointwise estimate for a quantity like  $|Du - G|$  (for any vector  $G \in \mathbb{R}^n$ ). Dimensionally speaking, this quantity represents an oscillation of the gradient when  $G$  is suitably chosen, i.e. when for instance is taken to be an average of the gradient. Then we conclude applying the sharp maximal inequality of Theorem 12 that in fact allows to estimate, in an integral way, oscillations of the gradient. Notice that in the rest of the proof we shall assume that

$$\mathbf{W}_{1 - \frac{(1+\alpha)(p-1)}{p}, p}^\mu(x, R) + \mathbf{W}_{1 - \frac{(1+\alpha)(p-1)}{p}, p}^\mu(y, R) < \infty \quad (199)$$

holds, otherwise there is nothing to prove. We notice that this condition allows to conclude that both  $x$  and  $y$  are Lebesgue points of the gradient  $Du$ ; this fact is indeed implied by Theorem 15, as the Wolff potentials appearing in (199) control the Riesz potential  $\mathbf{I}_1^\mu$  at the points  $x, y$  via a computation similar to the one in (26). We will anyway give a shorter proof of this fact to make the proof of Theorem 9 self-contained; see Lemma 9 below.

### 16.1 General setting

We readopt the general setting of Theorem 10. Therefore we start by  $x, y \in B_{R/4}$  and we select  $r < R/2$  to be determined later. Then we follow the scheme described in Sect. 11. In particular, we define the shrinking balls  $\{B_i\}$  as in (81); they are all centred at  $x$ , while eventually all the arguments will be reapplied when all the balls will be centred at  $y$ . The functions  $v_j$  are defined in (82) while as usual we denote

$$E_j := \int_{B_j} |Du - (Du)_{B_j}| dx, \quad j \geq 0.$$

The number  $\delta_1$  determining the shrinking rate of the balls  $B_j$  is instead defined as

$$\delta_1 := \left( \frac{1}{10^7 c_d} \right)^{1/\beta}, \quad (200)$$

where the constants  $c_d$  and  $\beta$  are from Theorem 17. This in turn yields a dependence of  $\delta_1$  on  $n, p, \nu, L$  only.

## 16.2 Estimates on the excess

By the choice of  $\delta_1$  we can apply Lemma 4 with  $\alpha = 1$ ; this gives

$$E_{j+1} \leq \frac{1}{10^6} E_j + 8c_1 c_d \delta_1^{-n} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \quad (201)$$

for every  $j \geq 0$ . By summing inequalities (201) for  $j \in \{k, \dots, h\}$ , where  $1 \leq k \leq h$  are two arbitrary nonnegative integers, we gain

$$\sum_{j=k+1}^{h+1} E_j \leq \frac{1}{10^6} \sum_{j=k}^h E_j + c \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)}.$$

so that

$$\sum_{j=k+1}^{h+1} E_j \leq E_k + c \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)}$$

and, ultimately

$$\sum_{j=k}^{h+1} E_j \leq 2E_k + c \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)},$$

where  $c \equiv c(n, p, \nu, L)$ . By noticing that

$$\left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} = r_j^{\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \leq r^{\alpha} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}$$

we conclude with

$$\sum_{j=k}^{h+1} E_j \leq 2E_k + cr^{\alpha} \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}. \quad (202)$$

Next, we estimate the last sum in terms of Wolff potentials as follows:

$$\begin{aligned}
 \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^{\mu}(x, 2r) &= \int_0^{2r} \left[ \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \frac{d\varrho}{\varrho} \\
 &= \sum_{j=0}^{\infty} \int_{r_{j+1}}^{r_j} \left[ \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \frac{d\varrho}{\varrho} \\
 &\quad + \int_r^{2r} \left[ \frac{|\mu|(B(x, \varrho))}{\varrho^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \frac{d\varrho}{\varrho} \\
 &\geq \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_{j+1})}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \int_{r_{j+1}}^{r_j} \frac{d\varrho}{\varrho} \\
 &\quad + \left[ \frac{|\mu|(B_0)}{(2r)^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \int_r^{2r} \frac{d\varrho}{\varrho} \\
 &\geq \delta_1^{(n-1)/(p-1)+\alpha} \log \left( \frac{1}{\delta_1} \right) \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_{j+1})}{r_{j+1}^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \\
 &\quad + \frac{\log 2}{2^{(n-1)/(p-1)+\alpha}} \left[ \frac{|\mu|(B_0)}{r_0^{n-1+\alpha(p-1)}} \right]^{1/(p-1)} \\
 &\geq \delta_1^n \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}. \tag{203}
 \end{aligned}$$

Now, by taking  $k = 0$  and  $h = i$  in (202) we have

$$\sum_{j=0}^{i+1} E_j \leq 2E_0 + cr^{\alpha} \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1+\alpha(p-1)}} \right]^{1/(p-1)}$$

and merging the content of the last two displays we conclude with

$$\sum_{j=0}^{i+1} E_j \leq 2E_0 + cr^{\alpha} \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^{\mu}(x, R) \tag{204}$$



where we have used the fact that  $2r \leq R$  and the constant  $c$  depends only on  $n, p, v, L$  thanks the choice of  $\delta_1$  made in (200). In particular, again taking  $\alpha = 0$  and eventually letting  $i \rightarrow \infty$  in the previous inequality gives

$$\sum_{j=0}^{\infty} E_j \leq 2E_0 + cW_{1/p,p}^{\mu}(x, R). \quad (205)$$

### 16.3 Pointwise estimate and Lebesgue point criterion

Let us prove

**Lemma 9** (Wolff potentials control Lebesgue points) *The condition*

$$W_{1/p,p}^{\mu}(x, R) < \infty \quad (206)$$

*implies that the following limit exists:*

$$\lim_{i \rightarrow \infty} (Du)_{B_i} = Du(x).$$

*Proof* Observe that exactly as in (140) we have, whenever  $h > k$

$$|(Du)_{B_h} - (Du)_{B_k}| \leq \delta_1^{-n} \sum_{j=k}^{h-1} E_j. \quad (207)$$

Using (202) with  $\alpha = 0$  we gain

$$|(Du)_{B_h} - (Du)_{B_k}| \leq cE_k + c \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)}. \quad (208)$$

Now, notice that (203) with  $\alpha = 0$  gives

$$\sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} \leq cW_{1/p,p}^{\mu}(x, 2r) < \infty \quad (209)$$

and this, together with (205) gives

$$\lim_{k \rightarrow \infty} E_k + \sum_{j=k}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-1}} \right]^{1/(p-1)} = 0.$$

In turn this last fact and (208) imply that  $\{(Du)_{B_i}\}$  is a Cauchy sequence and the assertion is proved. Notice that a slight variant of this argument (see for instance

Sect. 14.2) leads to establish that  $x$  is a Lebesgue point of  $Du$  in the usual sense. See also Sect. 17 below.  $\square$

As (199) implies the finiteness condition in (206) the previous lemma implies in particular that

$$\lim_{i \rightarrow \infty} |(Du)_{B_i} - G| = |Du(x) - G| \quad (210)$$

holds whenever  $G \in \mathbb{R}^n$ . Let us use (140) for the choice  $k_1 = 0$  and  $k_2 = i \geq 1$  to get

$$|(Du)_{B_i} - (Du)_{B_0}| \leq \delta_1^{-n} \sum_{j=0}^{i-1} E_j \equiv c \sum_{j=0}^{i-1} E_j$$

with  $c \equiv c(n, p, \nu, L)$ . Further use of triangle inequality then gives

$$|(Du)_{B_i} - G| \leq |(Du)_{B_0} - G| + c \sum_{j=0}^{i-1} E_j$$

so that, using (204) we obtain

$$|(Du)_{B_i} - G| \leq |(Du)_{B_0} - G| + cE_0 + cr^\alpha \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^\mu(x, R),$$

where the constant  $c$  depends only on  $n, p, \nu, L$ . Letting  $i \rightarrow \infty$  and recalling (210) we conclude with

$$\begin{aligned} |Du(x) - G| &\leq c \int_{B(x, r)} |Du - (Du)_{B(x, r)}| + |Du - G| d\tilde{x} \\ &\quad + cr^\alpha \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^\mu(x, R). \end{aligned}$$

#### 16.4 Use of the maximal estimate and conclusion

Writing inequality (211) for  $y$  instead of  $x$  and summing up gives

$$\begin{aligned} |Du(x) - Du(y)| &\leq c \int_{B(x, r)} (|Du - (Du)_{B(x, r)}| + |Du - G|) d\tilde{x} \\ &\quad + c \int_{B(y, r)} (|Du - (Du)_{B(y, r)}| + |Du - G|) d\tilde{x} \\ &\quad + c \left[ \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^\mu(x, R) + \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p}, p}^\mu(y, R) \right] r^\alpha. \end{aligned} \quad (211)$$

We recall that we can still choose  $G$  and  $r \leq R/2$ ; we therefore take  $G := (Du)_{B(x,3r)}$  and  $r := |x - y|/2$ , so that  $B(y, r) \subset B(x, 3r)$  and therefore, using also (98) we have

$$\begin{aligned} & \int_{B(x,r)} (|Du - (Du)_{B(x,r)}| + |Du - G|) d\tilde{x} \\ & + \int_{B(y,r)} (|Du - (Du)_{B(y,r)}| + |Du - G|) d\tilde{x} \\ & \leq c(n) \int_{B(x,3r)} |Du - (Du)_{B(x,3r)}| d\tilde{x}. \end{aligned} \quad (212)$$

Now, notice that as we are assuming that the initial ball  $B_R$  is such that  $x, y \in B_{R/4}$ , then we have necessarily  $|x - y| \leq R/2$  so that now  $r \leq R/4$  and  $B(x, 3r) \subset B(x, 3R/4) \subset B_R$ . Therefore we use the definition of sharp maximal operator and Theorem 12 to estimate

$$\begin{aligned} & \int_{B(x,3r)} |Du - (Du)_{B(x,3r)}| d\tilde{x} \leq cr^\alpha M_{\alpha,3r}^\#(Du)(x) \\ & \leq cr^\alpha [M_{1-\alpha(p-1),3R/4}(\mu)(x)]^{1/(p-1)} + c \left(\frac{r}{R}\right)^\alpha \int_{B_{3R/4}} |Du - (Du)_{B_{3R/4}}| d\tilde{x}. \end{aligned}$$

It is possible to estimate the fractional maximal operators by Wolff potentials via completely elementary means; for instance, the following inequality is proved in [40, Lemma 4.1] and holds whenever  $\gamma \in (0, 1)$ :

$$[M_{\beta,\gamma R}(\mu)(x)]^{1/(p-1)} \leq \frac{\max\{\gamma^{(\beta-n)/(p-1)}, 1\}}{(-\log \gamma)|B(0, 1)|^{1/(p-1)}} \mathbf{W}_{\beta/p,p}^\mu(x, R). \quad (213)$$

The content of the last two displays and (98) yields

$$\begin{aligned} & \int_{B(x,3r)} |Du - (Du)_{B(x,3r)}| d\tilde{x} \\ & \leq cr^\alpha \mathbf{W}_{1-\frac{(1+\alpha)(p-1)}{p},p}^\mu(x, R) + c \left(\frac{r}{R}\right)^\alpha \int_{B_R} |Du - (Du)_{B_R}| d\tilde{x}. \end{aligned}$$

Connecting the last inequality to (212) and eventually to (211) completes the proof of Theorem 9.

## 17 Proof of Theorem 16

The proof goes in two steps, reported in Sects. 17.1 and 17.2, respectively. We here give the whole proof of Theorem 16 in the case of an energy solution  $u \in W^{1,p}(\Omega)$ ;

the case of a general SOLA can be treated adapting the arguments valid for the energy case along the lines described in Sect. 14.4 for Theorem 15. In the rest of the proof all the balls will be centred at the point  $x$ , starting from an initial ball  $B(x, R) \subset \Omega$  with  $R \leq 1$ , which is not restrictive of course.

### 17.1 Smallness of the excess

Here we prove

$$\lim_{\varrho \rightarrow 0} E(u, B(x, \varrho)) = \lim_{\varrho \rightarrow 0} \int_{B(x, \varrho)} |u - (u)_{B(x, \varrho)}| d\tilde{x} = 0 \quad (214)$$

that is, we prove that for every  $\tilde{\varepsilon} > 0$  there exists  $\tilde{R} > 0$ , basically depending on  $\tilde{\varepsilon}$  and on the point  $x$ , such that

$$\varrho \leq \tilde{R} \implies E(u, B(x, \varrho)) < \tilde{\varepsilon}. \quad (215)$$

To start with, we notice that by using (44) with  $\alpha = 0$ , together with (213) with  $\beta = p$  to establish the finiteness of the relevant maximal operator, we get

$$\sup_{0 < \varrho \leq R} \int_{B(x, \varrho)} |u - (u)_{B(x, \varrho)}| d\tilde{x} < \infty. \quad (216)$$

Thanks to this we are able to determine  $\varepsilon > 0$  small enough in order to have

$$\varepsilon \sup_{0 < \varrho \leq R} \int_{B(x, \varrho)} |u - (u)_{B(x, \varrho)}| d\tilde{x} + 2\varepsilon R s \leq \frac{\tilde{\varepsilon}}{2}. \quad (217)$$

With  $\varepsilon$  been determined, we define  $\delta_1$  as

$$\delta_1 := \left( \frac{\varepsilon}{10^8 \tilde{c}_d} \right)^{1/\tilde{\beta}}, \quad (218)$$

where the constants  $\tilde{c}_d$  and  $\tilde{\beta}$  are from Theorem 18. In turn, having  $\delta_1$  being defined, we choose the radius  $r \leq R/2$  such that the smallness condition

$$16\tilde{c}_1 \tilde{c}_d \delta_1^{-3n} \mathbf{W}_{1,p}^\mu(x, 2r) \leq \frac{\tilde{\varepsilon}}{2} \quad (219)$$

is satisfied. The constant  $\tilde{c}_1$  has been introduced in (77). Notice that this is possible since the assumption in (62) and the definition of Wolff potential imply that

$$\lim_{\varrho \rightarrow 0} \mathbf{W}_{1,p}^\mu(x, \varrho) = 0. \quad (220)$$

Determining  $r$  allows to consider the sequence of shrinking balls  $\{B_j\}$  as defined in (81); the functions  $\{v_j\}$  are as in (112). Accordingly, we define

$$\tilde{E}_j := \int_{B_j} |u - (u)_{B_j}| d\tilde{x}.$$

Finally, an argument similar to the ones in (203) yields

$$\delta_1^n \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-p}} \right]^{1/(p-1)} \leq \mathbf{W}_{1,p}^{\mu}(x, 2r). \quad (221)$$

This is indeed an inequality that actually works whenever  $\delta_1 \in (0, 1/4)$  is a free parameter and the balls  $\{B_j\}$  are defined as  $B_j := B(x, \delta_1^j r)$ .

Now, the first thing we do is looking at Lemma 4 (when  $\alpha = 0$ ) and make computations similar to the one involved there, but using Theorem 18 instead of Theorem 17, and using (77) instead of (76). Specifically, for every  $j \geq 0$  we have

$$\begin{aligned} \tilde{E}_{j+1} &\leq 2 \int_{B_{j+1}} |v_j - (v_j)_{B_{j+1}}| d\tilde{x} + 2\delta_1^{-n} \int_{B_j} |v_j - u| d\tilde{x} \\ &\leq 2\tilde{c}_d \delta_1^{\tilde{\beta}} \int_{B_j} |v_j - (v_j)_{B_j}| d\tilde{x} + 2\tilde{c}_1 \delta_1^{-n} \left[ \frac{|\mu|(B_j)}{r_j^{n-p}} \right]^{1/(p-1)} + 2\tilde{c}_d \delta_1^{\tilde{\beta}} r_j s \\ &\leq 4\tilde{c}_d \delta_1^{\tilde{\beta}} \int_{B_j} |u - (u)_{B_j}| d\tilde{x} + 8\tilde{c}_1 \tilde{c}_d \delta_1^{-n} \left[ \frac{|\mu|(B_j)}{r_j^{n-p}} \right]^{1/(p-1)} + 2\tilde{c}_d \delta_1^{\tilde{\beta}} r_j s. \end{aligned}$$

Recalling (218) we conclude that the following estimate holds for every  $j \geq 0$ :

$$\tilde{E}_{j+1} \leq \frac{\varepsilon}{10^6} \tilde{E}_j + 8\tilde{c}_1 \tilde{c}_d \delta_1^{-n} \left[ \frac{|\mu|(B_j)}{r_j^{n-p}} \right]^{1/(p-1)} + \tilde{\varepsilon} r_j s. \quad (222)$$

Summing up the preceding inequalities and proceeding as in Sect. 16.2 yields

$$\sum_{j=k}^h \tilde{E}_j \leq \varepsilon \tilde{E}_{k-1} + \varepsilon r s \sum_{j=0}^{\infty} \delta_1^j + 16\tilde{c}_1 \tilde{c}_d \delta_1^{-n} \sum_{j=0}^{\infty} \left[ \frac{|\mu|(B_j)}{r_j^{n-p}} \right]^{1/(p-1)}, \quad (223)$$

which is valid whenever  $1 \leq k \leq h$  are integers. We then use (217) to estimate the first two terms in the right hand side of the previous inequality and (219)–(221) to estimate the remaining one; therefore we conclude with

$$\sum_{j=1}^{\infty} \tilde{E}_j \leq \tilde{\varepsilon}. \quad (224)$$

This last inequality in particular implies that

$$\int_{B(x, \varrho)} |u - (u)_{B(x, \varrho)}| d\tilde{x} \leq \tilde{\varepsilon} \quad \text{holds whenever } \varrho = r_j \text{ and } j \geq 1. \quad (225)$$

On the other hand, take now  $\varrho' \leq r_1 = \delta_1 r$  and determine  $k \geq 1$  such that  $r_{k+1} < \varrho \leq r_k$ . This means that  $\varrho' = \delta_1^k r'$  for  $r' \in (\delta_1 r, r)$ . Now, in the proof of (225) we replace  $r$  by  $r'$  and we get a new chain of cylinders  $B'_j = B(x, \delta^j r')$  to which we apply the same reasoning as above; then (225) follows for  $\varrho = r'_j$  and every  $j \geq 1$ . In particular it follows that  $E(u, B(x, \varrho')) \leq \tilde{\varepsilon}$  and this means now that (215) holds for  $\tilde{R} = \delta_1 r$  where  $r$  has been determined via the choice in (219).

## 17.2 Proof that $x$ is a Lebesgue point of $u$

We are now ready for the proof of (63), therefore showing that for every  $\varepsilon > 0$ , there exists a radius  $r_\varepsilon$ , depending on  $\varepsilon$ , such that

$$|(u)_{B(x, \varrho)} - (u)_{B(x, \tau)}| \leq \varepsilon \quad \text{holds whenever } 0 < \tau < \varrho \leq r_\varepsilon. \quad (226)$$

We now define

$$\delta_1 := \left( \frac{1}{10^8 \tilde{c}_d} \right)^{1/\tilde{\beta}} \quad (227)$$

where  $\tilde{c}_d$  appears in (106), and then we take a radius  $\tilde{R}_l$  such that

$$\sup_{0 < \varrho \leq \tilde{R}_l} E(u, B(x, \varrho)) + 16\tilde{c}_1 \tilde{c}_d \delta_1^{-2n} \mathbf{W}_{1,p}^\mu(x, 2\tilde{R}_l) + 2s\tilde{R}_l \leq \frac{\delta_1^n \varepsilon}{3}. \quad (228)$$

We notice that this is possible thanks to (214) and (220). With  $\delta_1$  and  $\tilde{R}_l$  having been defined, we then use the usual shrinking chain of balls  $\{B_j\}$  defined in (81) with  $r = \tilde{R}_l$ , so that, in particular we have  $r_j = \delta_1^j \tilde{R}_l$ ; again, the functions  $\{v_j\}$  are as in (112). Similarly to (207), whenever  $h > k \geq 1$  are integers it holds that

$$|(u)_{B_h} - (u)_{B_k}| \leq \delta_1^{-n} \sum_{j=k}^{h-1} \tilde{E}_j. \quad (229)$$

We now observe that (223) still works in this setting with  $\varepsilon = 1$ , according to the choice in (227). This is

$$\sum_{j=k}^h \tilde{E}_j \leq \tilde{E}_{k-1} + rs \sum_{j=0}^{\infty} \delta_1^j + 16\tilde{c}_1 \tilde{c}_d \delta_1^{-2n} \mathbf{W}_{1,p}^{\mu}(x, 2r),$$

and we also used (221), which is again an inequality that whenever  $\delta_1 \in (0, 1/4)$  is a fixed parameter. Combining the inequalities in the last two displays together with (228), we obtain that

$$|(u)_{B_h} - (u)_{B_k}| \leq \frac{\varepsilon}{3} \quad \text{holds for every choice of } h > k \geq 1. \quad (230)$$

We can now prove (226) with the choice  $r_{\varepsilon} := \delta_1 \tilde{R}_l$ . Indeed, let us fix  $0 < \tau < \varrho \leq r_{\varepsilon}$ . This means that there exist two integers  $k$  and  $h$ , such that  $1 \leq k \leq h$ ,

$$\delta_1^{k+1} \tilde{R}_l < \varrho \leq \delta_1^k \tilde{R}_l \quad \text{and} \quad \delta_1^{h+1} \tilde{R}_l < \tau \leq \delta_1^h \tilde{R}_l$$

hold. Applying (228) yields

$$\begin{aligned} |(u)_{B(x,\varrho)} - (u)_{B_{k+1}}| &\leq \int_{B_{k+1}} |u - (u)_{B(x,\varrho)}| d\tilde{x} \\ &\leq \frac{|B(x,\varrho)|}{|B_{k+1}|} \int_{B_{k+1}} |u - (u)_{B(x,\varrho)}| d\tilde{x} \\ &\leq \delta_1^{-n} E(Du, B(x,\varrho)) \leq \frac{\varepsilon}{3}, \end{aligned}$$

and, similarly,

$$|(u)_{B(x,\tau)} - (u)_{B_{h+1}}| \leq \frac{\varepsilon}{3}.$$

The last two inequalities and (230) establish (226) and the proof of Theorem 16 is complete.

## 18 Back to the roots: Theorem 1 (sketch)

The arguments developed for Theorem 16 allow to get a quick proof of Theorem 1, that we sketch in the following lines. We repeat the construction of Sects. 17.1–17.2 with  $\delta_1$  as in (227), but starting from any ball  $B(x, R) \subset \Omega$ , with  $r = R/2$ ; the sequence  $\{B_j\}$  is once again accordingly defined as in (81). In (229) we take  $k = 1$  and let  $h \rightarrow \infty$ ; this yields

$$|u(x) - (u)_{B_0}| \leq \delta_1^{-n} \sum_{j=1}^{\infty} \tilde{E}_j.$$

Using this last inequality together with (221) and (223) (here it is  $\varepsilon = 1$ ) yields

$$\begin{aligned} |u(x)| &\leq |(u)_{B_0}| + crs + \tilde{E}_0 + c\mathbf{W}_{1,p}^\mu(x, R) \\ &\leq c\mathbf{W}_{1,p}^\mu(x, R) + c \int_{B(x,R)} (|u| + Rs) d\tilde{x} \end{aligned}$$

for a constant  $c \equiv c(n, p, v, L)$ , that is (20). The proof of the continuity of  $u$  under assumption (21) follows instead making the arguments for the proof of Theorem 16 uniform with respect to  $x$ , in the same way we have done in Sect. 15 with respect to the proof of Theorem 15.

**Acknowledgments** We acknowledge the hospitality of the Universidad Autonoma de Madrid in April and May 2013, where part of this paper was written. In particular, we thank Matteo Bonforte and “Maestro” Juan Luis Vázquez for their kindness. The paper was finalised while we were both guests of the Institut Mittag-Leffler, joining the program “Evolutionary problems” in the fall 2013. Last but not least, we thank Paolo Baroni for precious remarks on a preliminary version of the paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Adams, D.R.: A note on Riesz potentials. *Duke Math. J.* **42**, 765–778 (1975)
2. Adams, D.R., Meyers, N.G.: Thinness and Wiener criteria for nonlinear potentials. *Indiana Univ. Math. J.* **22**, 169–197 (1972)
3. Adams, D.R., Hedberg, L.I.: *Function spaces and potential theory*. Grundlehren der Mathematischen Wissenschaften, vol. 314. Springer, Berlin (1996)
4. Aronsson, G., Crandall, M.G., Juutinen, P.: A tour of the theory of absolutely minimizing functions. *Bull. Am. Math. Soc. (N.S.)* **41**, 439–505 (2004)
5. Baroni, P.: Linear potential estimates for a class of elliptic equations (2013, preprint)
6. Baroni, P., Di Castro, A., Palatucci, G.: (in preparation)
7. Bénilan, P., Boccardo, L., Gallouët, T., Gariepy, R., Pierre, M., Vázquez, J.L.: An  $L^1$ -theory of existence and uniqueness of solutions of nonlinear elliptic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* **22**, 241–273 (1995)
8. Boccardo, L., Gallouët, T.: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* **87**, 149–169 (1989)
9. Byun, S.S., Wang, L., Zhou, S.: Nonlinear elliptic equations with BMO coefficients in Reifenberg domains. *J. Funct. Anal.* **250**, 167–196 (2007)
10. Campanato, S.: Proprietà di hölderianità di alcune classi di funzioni. *Ann. Scuola Norm. Sup. Pisa (III)* **17**, 175–188 (1963)
11. Cianchi, A.: Nonlinear potentials, local solutions to elliptic equations and rearrangements. *Ann. Scuola Norm. Sup. Cl. Sci. (V)* **10**, 335–361 (2011)
12. Cianchi, A., Maz’ya, V.: Global Lipschitz regularity for a class of quasilinear elliptic equations. *Commun. PDE* **36**, 100–133 (2011)
13. Dall’Aglio, A.: Approximated solutions of equations with  $L^1$  data. Application to the  $H$ -convergence of quasi-linear parabolic equations. *Ann. Mat. Pura Appl. (IV)* **170**, 207–240 (1996)
14. Dal Maso, G., Murat, F., Orsina, L., Prignet, A.: Renormalized solutions of elliptic equations with general measure data. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* **28**, 741–808 (1999)
15. Daskalopoulos, T., Kuusi, T., Mingione, G.: Borderline estimates for fully nonlinear elliptic equations. *Commun. PDE* (to appear)



16. De Giorgi, E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Nat. (III)* **125**(3), 25–43 (1957)
17. DeVore, R.A., Sharpley, R.C.: Maximal functions measuring smoothness. *Mem. Am. Math. Soc.* **47**, no. 293 (1984)
18. DiBenedetto, E.:  $C^{1+\alpha}$  local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal. TMA* **7**, 827–850 (1983)
19. DiBenedetto, E., Manfredi, J.J.: On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems. *Am. J. Math.* **115**, 1107–1134 (1993)
20. Dolzmann, G., Hungerbühler, N., Müller, S.: Uniqueness and maximal regularity for nonlinear elliptic systems of  $n$ -Laplace type with measure valued right hand side. *J. Reine Angew. Math.* **520**, 1–35 (2000)
21. Duzaar, F., Mingione, G.: Gradient estimates via non-linear potentials. *Am. J. Math.* **133**, 1093–1149 (2011)
22. Duzaar, F., Mingione, G.: Gradient estimates via linear and nonlinear potentials. *J. Funct. Anal.* **259**, 2961–2998 (2010)
23. Duzaar, F., Mingione, G.: Gradient continuity estimates. *Calc. Var. PDE* **39**, 379–418 (2010)
24. Evans, L.C.: A new proof of local  $C^{1,\alpha}$  regularity for solutions of certain degenerate elliptic p.d.e. *J. Differ. Equ.* **45**, 356–373 (1982)
25. Giusti, E.: *Direct Methods in the Calculus of Variations*. World Scientific Publishing Co., Inc, River Edge (2003)
26. Grafakos, L.: *Classical and Modern Fourier Analysis*. Pearson Edu. Inc., Upper Saddle River (2004)
27. Hamburger, C.: Regularity of differential forms minimizing degenerate elliptic functionals. *J. Reine Angew. Math. (Crelles J.)* **431**, 7–64 (1992)
28. Havin, M., Maz'ja, V.G.: Non-linear potential theory. *Russ. Math. Surveys* **27**, 71–148 (1972)
29. Hedberg, L., Wolff, Th.H.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* **33**, 161–187 (1983)
30. Heinonen, J., Kilpeläinen, T., Martio, O.: *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Oxford Mathematical Monographs, New York (1993)
31. Iwaniec, T.: Projections onto gradient fields and  $L^p$ -estimates for degenerated elliptic operators. *Studia Math.* **75**, 293–312 (1983)
32. Jaye, B., Verbitsky, I.: Local and global behaviour of solutions to nonlinear equations with natural growth terms. *Arch. Ration. Mech. Anal.* **204**, 627–681 (2012)
33. Jin, T., Mazya, V., Van Schaftingen, J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. *Comptes Rendus Mathématique* **347**, 773–778 (2009)
34. John, F., Nirenberg, L.: On functions of bounded mean oscillation. *Commun. Pure Appl. Math.* **14**, 415–426 (1961)
35. Kichenassamy, S., Verón, L.: Singular solutions to the  $p$ -Laplace equation. *Math. Ann.* **275**, 599–615 (1986)
36. Kilpeläinen, T., Kuusi, T., Tuhola-Kujanpää, A.: Superharmonic functions are locally renormalized solutions. *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **28**, 775–795 (2011)
37. Kilpeläinen, T., Malý, J.: Degenerate elliptic equations with measure data and nonlinear potentials. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* **19**, 591–613 (1992)
38. Kilpeläinen, T., Malý, J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* **172**, 137–161 (1994)
39. Korte, R., Kuusi, T.: A note on the Wolff potential estimate for solutions to elliptic equations involving measures. *Adv. Calc. Var.* **3**, 99–113 (2010)
40. Kuusi, T., Mingione, G.: Universal potential estimates. *J. Funct. Anal.* **262**, 4205–4269 (2012)
41. Kuusi, T., Mingione, G.: Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.* **207**, 215–246 (2013)
42. Kuusi, T., Mingione, G.: The Wolff gradient bound for degenerate parabolic equations. *J. Eur. Math. Soc.* (2014, to appear)
43. Kuusi, T., Mingione, G.: Gradient regularity for nonlinear parabolic equations. *Ann. Scu. Norm. Sup. Cl. Sci. (Ser. V)* **12**, 755–822 (2013)
44. Kuusi, T., Mingione, G.: Riesz potentials and nonlinear parabolic equations. *Arch. Ration. Mech. Anal.* doi:[10.1007/s00205-013-0695-8](https://doi.org/10.1007/s00205-013-0695-8)
45. Kuusi, T., Mingione, G.: A nonlinear Stein theorem. *Calc. Var. PDE*. doi:[10.1007/s00526-013-0666-9](https://doi.org/10.1007/s00526-013-0666-9)

46. Lewis, J.L.: Regularity of the derivatives of solutions to certain degenerate elliptic equations. *Indiana Univ. Math. J.* **32**, 849–858 (1983)
47. Lieberman, G.M.: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations. *Commun. PDE* **16**, 311–361 (1991)
48. Lindqvist, P.: On the definition and properties of  $p$ -superharmonic functions. *J. Reine Angew. Math. (Crelles J.)* **365**, 67–79 (1986)
49. Lindqvist, P.: Notes on the  $p$ -Laplace equation. Univ. Jyväskylä, Report 102 (2006)
50. Malý, J.: Wolff potential estimates of superminimizers of Orlicz type Dirichlet integrals. *Manuscr. Math.* **110**, 513–525 (2003)
51. Manfredi, J.J.: Regularity for minima of functionals with  $p$ -growth. *J. Differ. Equ.* **76**, 203–212 (1988)
52. Maz'ya, V., Shaposhnikova, T.: On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces. *J. Funct. Anal.* **195**, 230–238 (2002)
53. Mikkonen, P.: On the Wolff potential and quasilinear elliptic equations involving measures. *Ann. Acad. Sci. Fenn. Math. Diss.* **104** (1996)
54. Mingione, G.: The singular set of solutions to non-differentiable elliptic systems. *Arch. Ration. Mech. Anal.* **166**, 287–301 (2003)
55. Mingione, G.: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann. Scu. Norm. Sup. Pisa Cl. Sci. (V)* **6**, 195–261 (2007)
56. Mingione, G.: Gradient estimates below the duality exponent. *Math. Ann.* **346**, 571–627 (2010)
57. Mingione, G.: Gradient potential estimates. *J. Eur. Math. Soc.* **13**, 459–486 (2011)
58. Mingione, G.: Nonlinear measure data problems. *Milan J. Math.* **79**, 429–496 (2011)
59. Phuc, N.C., Verbitsky, I.E.: Quasilinear and Hessian equations of Lane–Emden type. *Ann. Math. (II)* **168**, 859–914 (2008)
60. Phuc, N.C., Verbitsky, I.E.: Singular quasilinear and Hessian equations and inequalities. *J. Funct. Anal.* **256**, 1875–1906 (2009)
61. Serrin, J.: Pathological solutions of elliptic differential equations. *Ann. Scuola Norm. Sup. Pisa (III)* **18**, 385–387 (1964)
62. Serrin, J.: Local behavior of solutions of quasi-linear equations. *Acta Math.* **111**, 247–302 (1964)
63. Stein, E.M.: Editor's note: the differentiability of functions in  $\mathbb{R}^n$ . *Ann. Math. (II)* **113**, 383–385 (1981)
64. Teixeira, E.: Sharp regularity for general Poisson equations with borderline sources. *J. Math. Pures Appl.* **99**, 150–164 (2013)
65. Trudinger, N.S., Wang, X.J.: On the weak continuity of elliptic operators and applications to potential theory. *Am. J. Math.* **124**, 369–410 (2002)
66. Trudinger, N.S., Wang, X.J.: Quasilinear elliptic equations with signed measure data. *Disc. Contin. Dyn. Syst. A* **23**, 477–494 (2009)
67. Ural'tseva, N.N.: Degenerate quasilinear elliptic systems. *Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **7**, 184–222 (1968)